



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

KEELE UNIVERSITY

**Investigation of the pituitary epigenome: a genome-wide
analysis of changes associated with sporadic tumours**

Cuong V. Duong

A dissertation submitted for the degree of Doctor of Philosophy
at the Keele University

February 2014

KEELE UNIVERSITY

**Investigation of the pituitary epigenome: a genome-wide
analysis of changes associated with sporadic tumours**

Student: Cuong V. Duong

Supervisor: William Farrell

A dissertation submitted for the degree of Doctor of Philosophy
at the Keele University

February 2014

Abstract

Pituitary tumours harbour epigenetic aberrations; however, characterisation of these aberrations, on a genome-wide basis, is hampered by their infrequent occurrence and their small size. To overcome the constraint of limited tissue, whole genome amplification of sodium bisulphite converted DNA was employed and provided a consistent 25-fold amplification from individual samples. This material was used in a genome-wide analysis the DNA methylation of 27,578 CpG sites for each of the major adenoma subtypes. In a discovery cohort, on the basis of stringent criteria, pyrosequencing validated 12 of 16 hypermethylated genes. Overall, the criteria identified 40 genes in non-functional, 21 in growth hormone, six in prolactin and two in corticotrophinoma. In an independent cohort, different frequencies of hypermethylation were apparent for each of these genes; however, association between methylation and reduced transcript expression was infrequent.

For the *EFEMP1* gene, following its initial identification, studies of an independent cohort of tumours showed frequent reduced *EFEMP1* expression, irrespective of adenoma subtype. However, reduced expression was not invariantly associated CpG island methylation. Conversely, chromatin immunoprecipitation assays (ChIP) showed histone modifications that were consonant with expression status. The causal relationship between gene silencing and epigenetic change was established by observing that epidrug challenges induced re-expression of *EFEMP1* in pituitary cells that was concomitant with histone modification associated with expressed genes.

Enforced expression of *EFEMP1* was without effect on cell proliferation or apoptotic end-points but was responsible for decreased expression of the *MMP2* transcript. This association was not apparent in primary adenomas, however, *MMP7*, showed a positive correlation with *EFEMP1* and this may reflect cell or species specific differences,

suggesting that the relationship between *EFEMP1* and *MMP7* requires more detailed investigation. This study is the first whole genome identification of a potential biomarker signature and their functional characterisation will provide insight of tumour aetiology and identification of new therapeutic targets.

Contents

<i>Tables</i>	<i>viii</i>
<i>Figures</i>	<i>ix</i>
<i>Abbreviations</i>	<i>x</i>
<i>Acknowledgements</i>	<i>xii</i>
<i>Publications</i>	<i>xiii</i>
<i>Section A: Introduction</i>	<i>1</i>
Chapter 1: Introduction	2
1.1. Epigenetics	2
1.1.1. DNA methylation	3
I. CpG islands	4
II. Heritability of DNA methylation	5
III. DNA demethylation	8
IV. Abnormal regulation of DNA methylation in cancer	9
IV.1. Global hypomethylation	10
IV.2. Gene specific CpG island hyper- and hypo-methylation	11
IV.3. Aberrant methylation of CpG island shores	13
IV.4. Abnormal imprinting	14
1.1.2. DNA hydroxymethylation	15
1.1.3. Histone modification	17
1.1.4. Mechanistic interdependence between DNA methylation and histone modifications	20
1.1.5. Pharmacological reversal of epigenetically silenced gene	22
I. DNMT inhibitors	23
II. HDAC inhibitors	23
III. Combinations of DNMT and HDAC inhibitors	24
IV. Combinations with conventional therapies	25
V. Genome-wide and gene-specific reactivation of silenced genes	25
V.1. Genome wide studies	25
V.2. Pharmacologic unmasking for gene specific investigations	26
1.2. The pituitary gland	28
1.2.1. Pituitary adenomas	28
1.2.2. Epigenetic changes in pituitary tumour cells	30
I. Cell-cycle regulators	30
II. Signal transduction pathways	31
III. Hormone and growth factor receptors	32
IV. Transcriptional regulators	33
1.3. Identification of novel genes involved in pituitary tumorigenesis	36
1.3.1. Limitation and advances in studies of epigenetic changes associated with pituitary adenomas	36
I. Candidate gene approach	36
II. Subtype specific	37
III. Species boundaries	38
IV. Limitation of clinical sample	38
	iv

1.3.2. DNA methylation profiling strategies	39
<i>I. Underlying methodologies for detection of DNA methylation</i>	40
I.1. Sodium bisulphite conversion	40
I.2. Methylation sensitive restriction enzymes	40
I.3. Methylated DNA immunoprecipitation	41
<i>II. Genome-wide DNA methylation profiling strategies</i>	42
II.1. Classical strategies	46
II.2. Array hybridization based strategies	48
II.3 NGS-based strategies	52
1.3.3. DNA hydroxymethylation profiling strategies	55
<i>I. Glucosylation methods</i>	56
<i>II. Affinity enrichment methods</i>	56
1.4. Aims of this study	58
Section B: Materials and Methods	60
Chapter 2: Material and methods	61
2.1. Primary Tissue material	61
2.1.1. Normal Pituitary Tissue	61
2.1.2. Pituitary Tumours	61
2.2. Cell culture methods	62
2.2.1. Pituitary Cell Lines:	62
<i>I. AtT20 Cells</i>	62
<i>II. GH3 Cells</i>	63
2.2.2. Growth Conditions	63
2.2.3. Antibiotic Supplementation	63
2.2.4. Detachment of adherent cells	64
2.2.5. Culture Vessels	64
2.2.6. Sub-culture	64
2.2.7. Cryopreservation of cells	65
2.2.8. Reviving cryopreserved cells	65
2.2.9. Cell Counting	66
2.2.10. Incubation of cell lines with pharmacological agents	66
2.3. Molecular biology methods	67
2.3.1. Isolation of genomic DNA from cell line and primary pituitary tissue	67
2.3.2. RNA extraction from cell lines and primary pituitary tissue	68
2.3.3. Purification of nucleic acid	69
2.3.4. Quantification of nucleic acid and assessment of purity	69
2.3.5. Agarose gel electrophoresis	70
2.4. Whole genome amplification of bisulphite converted DNA	71
2.4.1. Sodium bisulphite modification of genomic DNA	71
2.4.2. Whole genome amplification	72
2.5. DNA methylation analysis	73
2.5.1. Genome-wide DNA methylation analysis by Illumina BeadArray assay	73
<i>I. Sample preparation and processing</i>	73
<i>II. Data analysis</i>	73
2.5.2. Bisulphite sequencing	74
<i>I. Primer design for sodium bisulphite converted DNA</i>	74
<i>II. PCR amplification</i>	75
<i>III. Purification of DNA from agarose gels</i>	75

IV. Ligation	76
V. Transformation	77
VI. Spread plating	78
VII. Screening	78
2.5.3. Pyrosequencing	79
I. Assay design	81
II. PCR amplification	81
III. Pyrosequencing	82
2.6. Gene expression analysis	83
2.6.1. cDNA synthesis	83
2.6.2. Primer design	84
2.6.3. qRT-PCR	84
2.7. Histone modification analysis	86
2.7.1. Cell collection and DNA-protein crosslinking	88
2.7.2. Cell lysis and enzymatic chromatin shearing	89
2.7.3. Immunoprecipitation	90
2.7.4. Reversal of Crosslinking and DNA isolation.	91
2.7.5. Quantitative PCR and data analysis.	92
2.8. Functional analysis	94
2.8.1. Subcloning of target gene into pTUNE vector	94
2.8.2. Stable transfection of rat pituitary cell line GH3	95
2.8.3. Characterization of stable transfected cells	96
2.8.4. Cell proliferation analysis by growth curve	96
2.8.5. Apoptosis analysis by caspase 3/7 assay	96
2.8.6. Cell correction using crystal violet staining	98
Section C: Results and Discussions	99
Chapter 3: Whole genome amplification of bisulphite converted DNA	100
3.1. Whole genome amplification of bisulphite converted DNA	100
3.2. Integrity of methylation profile of BSC DNA prior to and following WGA	104
3.2.1. Sanger sequencing	104
3.2.2. Pyrosequencing	105
3.3. Discussion	106
Chapter 4: Genome-wide DNA methylation profiling of sporadic pituitary tumours	110
4.1. Methylation profiling of sporadic pituitary adenoma	110
4.2. Technical validation of Bead array by pyrosequencing	112
4.3. CpG island methylation profiles in primary adenoma subtypes	114
4.4. Gene specific methylation frequencies in non-functioning adenomas	120
4.5. Subtype specific gene methylation frequencies	123
4.6. Methylation associated gene silencing	125
4.7. Gene ontology analysis of hypermethylated genes	127
4.8. Discussion	128
Chapter 5: The <i>EFEMP1</i> gene: a frequent target for epigenetic silencing in multiple human pituitary adenoma subtypes	133
5.1. Expression of <i>EFEMP1</i> in pituitary adenomas	134
5.2. Methylation status of the <i>EFEMP1</i> genes CpG island in pituitary adenomas	136
5.3. Histone modifications associated with <i>EFEMP1</i> expression in primary adenoma	138
5.4. Epidrug induced re-expression of the <i>Efemp1</i> gene in AtT20 and GH3 cells	140

5.5. Epigenomic changes associated with expression of <i>Efemp1</i> in pituitary cell lines	142
5.6. Histone modifications in epidrug challenged pituitary cell lines	144
5.7. Function end-points contingent on heterologous <i>EFEMP1</i> expression in GH3 cells	146
5.7.1. Proliferation and apoptotic end-points	146
5.7.2. <i>EFEMP1</i> target gene end-points	146
5.8. Discussion	149
Chapter 6: Conclusions and future works	155
6.1. Conclusions	155
6.2. Suggestions for future work	159
6.2.1. Up-to-date genome-wide DNA methylation profiling of pituitary tumours	159
6.2.2. Further investigation of the candidate genes identified in this thesis	160
7.2.3. Further functional studies of <i>EFEMP1</i>	161
Section E: Appendices	163
Appendix 1: Primary pituitary adenomas used within the study	164
Appendix 2: Commonly used solutions	166
2.1. DNA extraction	166
2.2. RNA extraction	166
2.3. cDNA synthesis	167
2.4. Gel electrophoresis	167
2.5. Bacterial transformation	168
Appendix 3: Primer sequences	170
3.1. Bisulphite sequencing primers	170
3.2. Pyrosequencing primers	171
3.3. qRT-PCR primers	176
3.4. ChIP primers	177
Appendix 4: Vector maps	178
Appendix 5: List of hypermethylated genes fulfil criteria of 0.4 in a single CpG site -	180
5.1. - within the CpG island	180
5.2. - outside of a CpG island	186
Appendix 6: Gene ontology of hypermethylated genes	190
6.1. Go biological process	190
6.2. Go molecular function	204
6.3. Kegg pathway	208
Appendix 7: Quality control of 15 pituitary samples in Illumina Infinium 27 BeadArray	209
Appendix 8: Expression of <i>CDKN2A</i> in 13 NFs investigation cohort relative to 4 NPs	210
References	211

Tables

Table 1.1. Types of pituitary adenomas.	29
Table 1.2. Epigenetic changes associated with pituitary tumours.	35
Table 1.3.1. Genome-wide DNA methylation analysis: sample pretreatments	43
Table 1.3.2. Genome-wide DNA methylation analysis: readout methods.....	44
Table 2.1. Ligation Reaction. Ligation reagents for one reaction.	77
Table 4.1. Identity of 22 randomly selected genes from the BeadArray dataset.	115
Table 4.2. Genes identified on the BeadArray as hypermethylated.	118
Table 4.3. Frequencies of gene-specific methylation in an independent cohort of NF adenomas.	121
Table 5.1. EFEMP1 expression status in pituitary adenoma.	136
Table 5.2. EFEMP1 expression status in pituitary adenoma relative to CpG island methylation status.....	137

Figures

<i>Figure 1.1. Conversion of cytosine to 5-methylcytosine.....</i>	<i>3</i>
<i>Figure 1.2. DNA methylation profile in normal cells in mammals.</i>	<i>5</i>
<i>Figure 1.3. Establishment and maintenance of DNA methylation in mammalian cells.</i>	<i>7</i>
<i>Figure 1.4. Abnormal DNA methylation in cancer.</i>	<i>10</i>
<i>Figure 1.5. Similarities and differences in cancer-associated hypo- and hyper-methylation of DNA.</i>	<i>13</i>
<i>Figure 1.6. Epigenetic modifications associated with transcriptionally active and transcriptionally silent genes.</i>	<i>19</i>
<i>Figure 1.7. Interdependence of DNA methylation and histone modifications.....</i>	<i>21</i>
<i>Figure 1.8. Sample throughput versus genome coverage.</i>	<i>45</i>
<i>Figure 1.9. Principle of Illumina Infinium Methylation assay.</i>	<i>52</i>
 <i>Figure 2.1. Representative pyrosequencing pyrogram.</i>	 <i>80</i>
<i>Figure 2.2. Schematic of chromatin Immunoprecipitation using ChIP-ITTM Express.....</i>	<i>88</i>
<i>Figure 2.3. Gel analysis of optimal enzymatic shearing.</i>	<i>90</i>
<i>Figure 2.4. ChIP calculations.</i>	<i>93</i>
 <i>Figure 3.1. Whole genome amplification results in amplification of bisulphite converted DNA</i>	 <i>101</i>
<i>Figure 3.2. Representative examples of sequence-specific nested PCRs using dilution series of WGA- and WGA+ as templates.....</i>	<i>103</i>
<i>Figure 3.3. Sanger sequencing of MBD3 and GLRX.</i>	<i>105</i>
<i>Figure 3.4. Correlation of methylation status produced by pyrosequencing derived from WGA- and WGA+.</i>	<i>106</i>
 <i>Figure 4 1. Correlation between BeadArray-derived β values (expressed as a percentage) and pyrosequencing across a range on β values from 12 adenomas for four CpGs</i>	 <i>113</i>
<i>Figure 4.2. Heatmap of genes in individual NF and GH adenomas that fulfilled the criteria of elevated β in two gene-associated CpGs relative to the NP.</i>	<i>119</i>
<i>Figure 4.3. Pyrosequence analysis of promoter-associated CpG island methylation of genes first identified as hypermethylated by BeadArray analysis.</i>	<i>122</i>
<i>Figure 4.4. Adenoma subtype specificity of hypermethylation for three genes in which pyrosequencing confirmed their methylation status in NF adenomas is shown in Figure 4.3.</i>	<i>124</i>
<i>Figure 4.5. Quantitative RT-PCR analysis for three of 12 gene transcripts that show an inverse relationship between hypermethylation and decrease in transcript expression.</i>	<i>126</i>
 <i>Figure 5.1. Quantitative RT-PCR analysis of EFEMP1 expression in each of the major pituitary adenoma subtypes.</i>	 <i>135</i>
<i>Figure 5.2. ChIP Analysis of the EFEMP1 gene.....</i>	<i>139</i>
<i>Figure 5.3. Expression status of Efemp1 in pituitary cell lines.....</i>	<i>141</i>
<i>Figure 5.4. Inverse correlation between transcriptional expression and promoter methylation of Efemp1 in rat and mouse pituitary models.</i>	<i>143</i>
<i>Figure 5.5. ChIP analysis of Efemp1 promoter region in pituitary cell lines.</i>	<i>145</i>
<i>Figure 5.6. MMP2 expression in stable transfected GH3 cells.....</i>	<i>147</i>
<i>Figure 5.7. Expression of MMP7 in pituitary adenoma in the absence and presence of EFEMP1.</i>	<i>148</i>

Abbreviations

ACTH	Adrenocorticotroph hormone
ATP	Adenosine tri-phosphate
cDNA	Complementary DNA
CGI	CpG island
ChIP	Chromatin immunoprecipitation
CpG	Cytosine preceding Guanine
DMEM	Dulbecco's modified Eagles medium
DMR	Differently methylated region
DMSO	Dimethyl sulphoxide
DNA	Deoxyribo Nucleic Acid
DNMT	DNA methyl transferase
dNTP	Deoxynucleoside triphosphate
E. coli	Escherichia coli
EDTA	Ethylene Diamine Tetraacetic Acid
GH	Growth Hormone
H ₂ O	Water
H3K27me ₃	Histone 3 lysine 27 trimethylation
H3K9Ac	Histone 3 lysine 9 acetylation
HCl	Hydrochloric acid
HDAC	Histone deacetylase
HMT	Histone methyltransferase
IPTG	Isopropyl B-D-thiogalactopyranoside
LB	Luria Broth
LH	Luteinizing hormone
LOH	Loss of heterozygosity
LOI	Loss of imprinting
MBD	Methyl binding protein
MgCl ₂	Magnesium chloride
M-MLV	Moloney murine leukemia virus
mRNA	messenger RNA
MsAP PCR	Methylation specific PCR
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NF	Non functioning
NMP	Normal mouse pituitary
NP	Normal pituitary
NRP	Normal rat pituitary
NSCLC	Non small cell lung cancer
OD	Optical density
PBS	Phosphate buffered serum

PCR	Polymerase chain reaction
PMSF	Phenylmethanesulphonylfluoride
Ppi	Pyrophosphate
PRL	Prolactin hormone
PTAG	Pituitary tumour apoptosis gene
qRT-PCR	Quantitative reverse transcriptase PCR
RASSF1	Ras association domain family 1A gene
RLGS	Restriction landmark genome scanning
RNA	Ribonucleic acid
RT-PCR	Reverse transcription PCR
SAM	S-adenosyl-methionine
Taq	Thermo aquaticus
TSG	Tumour suppressor gene
X-gal	5-bromo-4-chloro-3-indolyl B-D-galactopyranonide
5-mC	5-methylcytosine
5-hmC	5-hydroxymethylcytosine
LINE-1	Long interspersed nuclear element-1
SINEs	Short interspersed nuclear elements
BSC	Sodium bisulphite conversion
WGA	Whole genome amplification
HAT	Histone acetyl transferase
TF	Transcription factor
Pol II	RNA polymerase II
LSD	Lysine demethylases
Zebularine	1-(B-D-ribofuranosyl)-1,2-dihydropyrimidin-2-one
TSA	Trichostatin A
siRNA	Small interfering RNA
FSH	Follicle-stimulating hormone
LH	Luteinizing hormone
TSH	Thyroid-stimulating hormone
COBRA	Combined bisulphite restriction analysis
AIMS	Amplification of inter-methylated sites
BS	Bisulphite sequencing
DMH	Differential methylation hybridization
HELP	HpaII tiny fragment enrichment by ligation-mediated PCR
MCA	Methylated CpG island amplification
MeDIP	Methylated DNA immunoprecipitation
MIRA	Methylated CpG island recovery assay
NGS	next-generation sequencing
RRBS	Reduced representation bisulphite sequencing
WGSBS	Whole-genome shotgun bisulphite sequencing
mDIP or mCIP	Methylcytosine immunoprecipitation

Acknowledgements

Foremost, I would like to express my deepest gratitude to my supervisor Professor William E. Farrell for his valuable support of my Ph.D study. He has always been supportive, patient, disciplined, and enthusiastic in all aspects of my study and research. I am grateful for his scientific advices, insightful comments and weekly discussions. It is my lucky to have such a brilliant supervisor.

I would like to express my sincere gratitude to Vietnam International Education Development - Ministry of Education and Training for funding my study.

I thank Dr. Alan Richardson for his encouragement, insightful comments, and hard questions. I also would like to thank Dr. Richard Emes and Frank Wessely for their collaboration in the genome-wide DNA methylation research.

I thank my lab-mates: Dr. Kim Haworth, Kiren Yacqub Usman, and Mark Kitchen for their kindness, friendships and supports. I will never forget their valuable helps when I first come to the UK. They are not only lab-mates but also my closest friends who are always ready to discuss scientific experiments as well as friendly talks.

Last but not the least, I would like to thank my family for all of the best things they provided. I thank my wife Hue for her care and patience at all times. Thank you my boy Minh Duy for being our great son during one of the most emotional periods of our lives.

Publications

Duong, C. V., Emes, R. D., Wessely, F., Yacqub-Usman, K., Clayton, R. N. & Farrell, W. E. (2012). *Quantitative, genome-wide analysis of the DNA methylome in sporadic pituitary adenomas*. **Endocr Relat Cancer** 19(6): 805-816.

Yacqub-Usman, K., Duong, C. V., Clayton, R. N. & Farrell, W. E. (2012). *Epigenomic silencing of the BMP-4 gene in pituitary adenomas: a potential target for epidrug-induced re-expression*. **Endocrinology** 153(8): 3603-3612.

Yacqub-Usman, K., Richardson, A., Duong, C. V., Clayton, R. N. & Farrell, W. E. (2012). *The pituitary tumour epigenome: aberrations and prospects for targeted therapy*. **Nat Rev Endocrinol** 8(8): 486-494.

Yacqub-Usman, K., Duong, C. V., Clayton, R. N. & Farrell, W. E. (2013). *Pre-incubation of pituitary tumour cells with the epidrugs zebularine and trichostatin A is permissive for retinoic acid augmented expression of the BMP-4 and D2R genes*. **Endocrinology** 154(5): 1711-1721.

Duong, C. V., Yacqub-Usman, K., Emes, R. D., Clayton, R. N. & Farrell, W. E. *The EFEMP1 gene: a frequent target for epigenetic silencing in multiple human pituitary adenoma subtypes*. *Neuroendocrinology*, 2013.

Section A: Introduction

Chapter 1: Introduction

1.1. Epigenetics

The genetic information within a genome is essentially invariant in each of the different cell types of an organism. However, between different cell types, despite the fact that they harbour the same invariant genetic information, different gene expression patterns are apparent prior to and following differentiation. An explanation for these findings is provided through an additional layer of heritable information termed “epigenetics”. The current accepted definition of epigenetics is ‘a process that heritably influences the expression of a gene without genetic change to the underlying DNA sequence itself.’ (Jaenisch and Bird, 2003). In these cases, heritability is principally mediated by post-replicative modifications that include methylation of cytosine bases in DNA and modification(s) of histone proteins and expression of micro RNAs (Jones and Baylin, 2007). The combinatorial influences of these changes frequently lead to chromatin remodelling or effects on translation. The resultant remodelling regulates access to the underlying genetic information by transcription factors and can result in inappropriate activation or inhibition of genes. Moreover, the expression status of these epigenetically modified genes frequently impacts upon down-stream signalling pathways and lead to disease states including cancer (Egger *et al.*, 2004; Jones and Baylin, 2002). This chapter will provide an overview of the principle epigenetic modification in healthy cells and those apparent in tumours. Furthermore, it will focus attention on the fact that epigenetic aberrations, unlike genetic mutations, are potentially reversible and their restoration to the normal state may be achieved through epigenetic therapies. Particular attention is given to specific reversal/restoration strategies and their impact on our understanding of tumour biology and on the clinical management of tumours including those of pituitary origin.

1.1.1. DNA methylation

DNA methylation was first described from the tubercle bacillus (Johnson and Coghill, 1925) and then in calf thymus (Hotchkiss, 1948). It is the modification of nucleotides in which a methyl group is covalently added to the 5-carbon position of cytosine residue (Figure 1) in the DNA double helix to form 5-methylcytosine (5-mC) (Wyatt, 1951). The DNA methylation reaction is catalyzed by methyltransferase with S-adenosyl serving as methyl donor results in S-adenosylmethionine and 5-mC. Because of its frequency and importance, 5-mC has been called the fifth nucleotide (Doerfler, 2006).

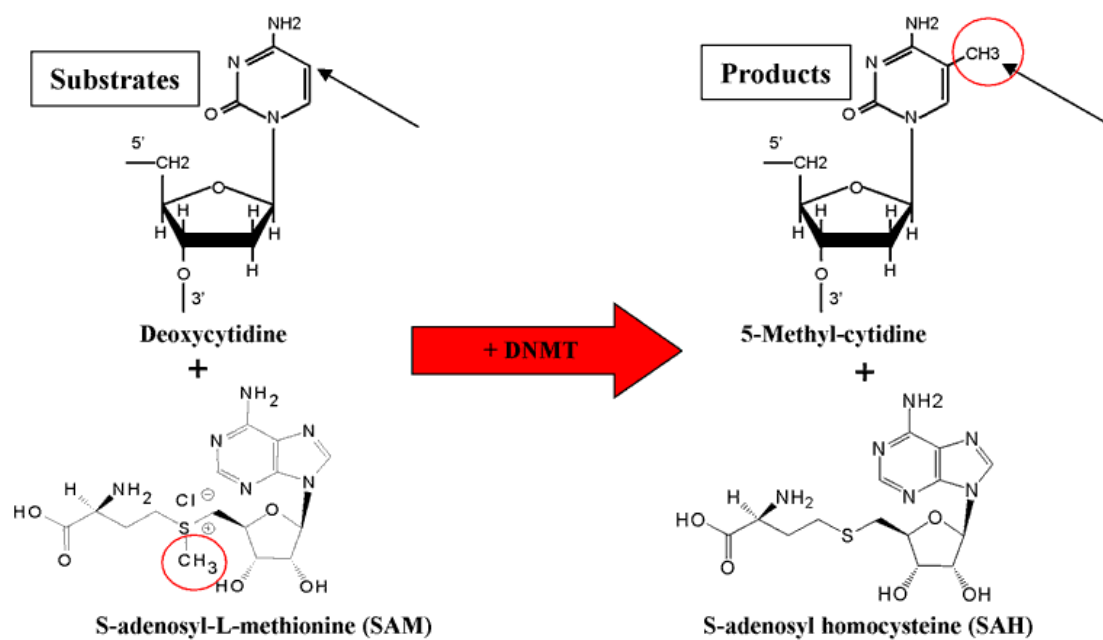


Figure 1.1. Conversion of cytosine to 5-methylcytosine

The arrows point to the 5-carbon position before and after methylation (CH₃). DNA methyltransferases is shown as, DNMT.

In eukaryotes, the majority of DNA methylation is found at cytosine residues in the context of cytosine-phosphate-guanine (CpG) dinucleotides (Gorovsky *et al.*, 1973;

Cummings *et al.*, 1974; Hattman *et al.*, 1978). Approximately 70% to 80% of CpG dinucleotides, throughout the genome, in human somatic cells are associated with methylation (Ehrlich *et al.*, 1982) whereas around 1% of all cytosine, that are not associated with a G residue, are thought to be methylated (Kriaucionis and Bird, 2003). In mammalian cells, the majority of 5-mC occurs within repetitive DNA sequence (Yu *et al.*, 2001). A large portion of these are located in transposons, such as *Alu* elements and L1 sequences (Smit and Riggs, 1996). The literature suggests that more than 45% of the human genome is represented by transposons (Lander *et al.*, 2001).

I. CpG islands

The overall frequency of CpG dinucleotides across the mammalian genome is approximately 1%, however, in some regions of the genome the observed frequency and ratio of CpG dinucleotides is higher at 4-6% (Kriaucionis and Bird, 2003). As previously discussed the majority of CpG dinucleotides in the genome are methylated whereas most of CpG sites in these “CpG enriched regions” are not (Figure 1.2). These regions are referred to as CpG islands (CGIs) and are DNA region which contains high frequency of CpG sites. In 1987, Gardiner-Garden proposed that a CGI should not be smaller than 200 bp with CG percentage greater than 50% and CpG ratio greater than 60% (Gardiner-Garden and Frommer, 1987). In 2002, by looking intensively at human chromosome 21 and 22, Takai and Jones showed a different value for CGIs where the smallest length, percentage of CG, and CpG ratio is 500 bp, 55%, and 65%, respectively (Takai and Jones, 2002).

CGIs are frequently found to be associated with the promoter regions of genes (Feil and Berger, 2007). In the majority of these cases, the default setting, with the exception of X-chromosome inactivation, the imprinted genes, LINE-1/SINES and some tissue specific

genes, are that these CGIs are not methylated. The promoter regions of approximately half of all transcribed genes are associated with CGIs. In general, the islands are found to encompass the promoter, or a proportion thereof and the first exon, however, several islands are also observed at a distance and/or downstream of the promoter region (Jones, 1986). It has been suggested that the association between gene silencing is more frequent when the CpG island is found within promoter region and extending into the first coding exon (Ushijima, 2005).

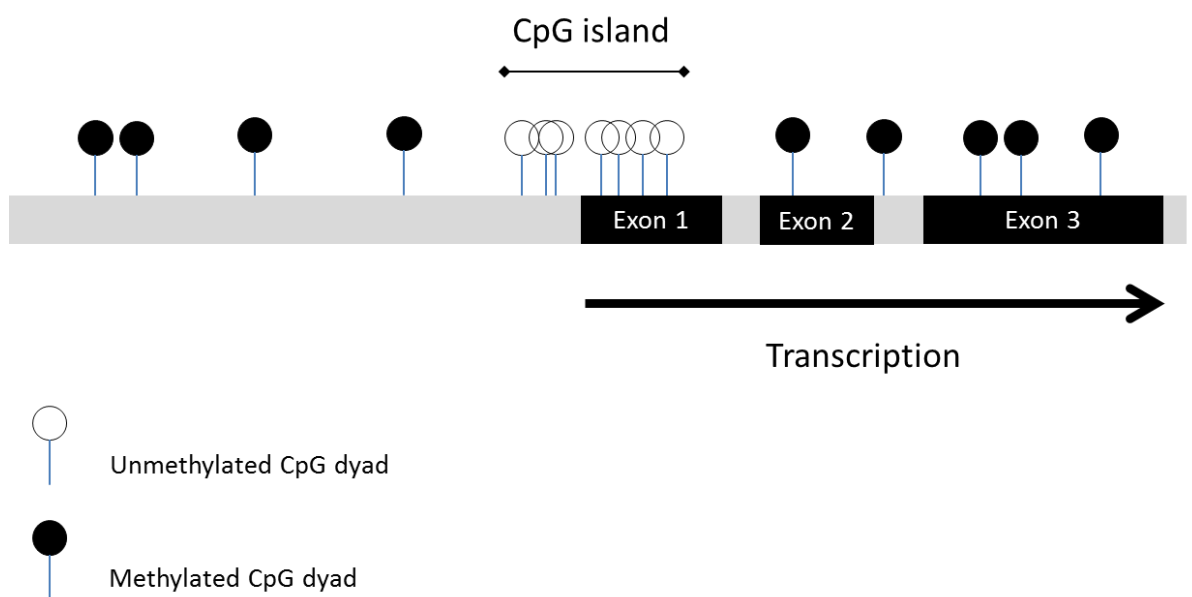


Figure 1.2. DNA methylation profile in normal cells in mammals.

CGIs are not, in general, methylated and non-CGI CpGs are frequently methylated.

II. Heritability of DNA methylation

DNA methylation reactions are catalysed by a conserved group of eukaryote proteins named DNA (cytosine-5) methyltransferases (DNMTs). DNMT methylate cytosine at carbon 5 (also see figure 1.1) resulted in 5-methylcytosine (Klimasauskas *et al.*, 1994). DNA methylation is a post synthetic modification where 5-mC is established after DNA

replication. Once formed, the modification is stable and can be transferred to the next generation during cell division (Reik and Dean, 2001). For example, the silenced *Lmyc* gene was found to be heritable not because of nucleotide sequence change but due instead to inherited DNA methylation patterns (Cubas *et al.*, 1999).

There are two principle mechanisms that are thought to be responsible for DNA methylation (Riggs, 1975; Holliday and Pugh, 1975). The first is *de novo* methylation, that creates a DNA methylation profile during the early stage of development. The second is maintenance methylation, and as its name implies is responsible for the inheritance of previously formed DNA methylation patterns during cell division and to the daughter cell(s) (Figure 1.3).

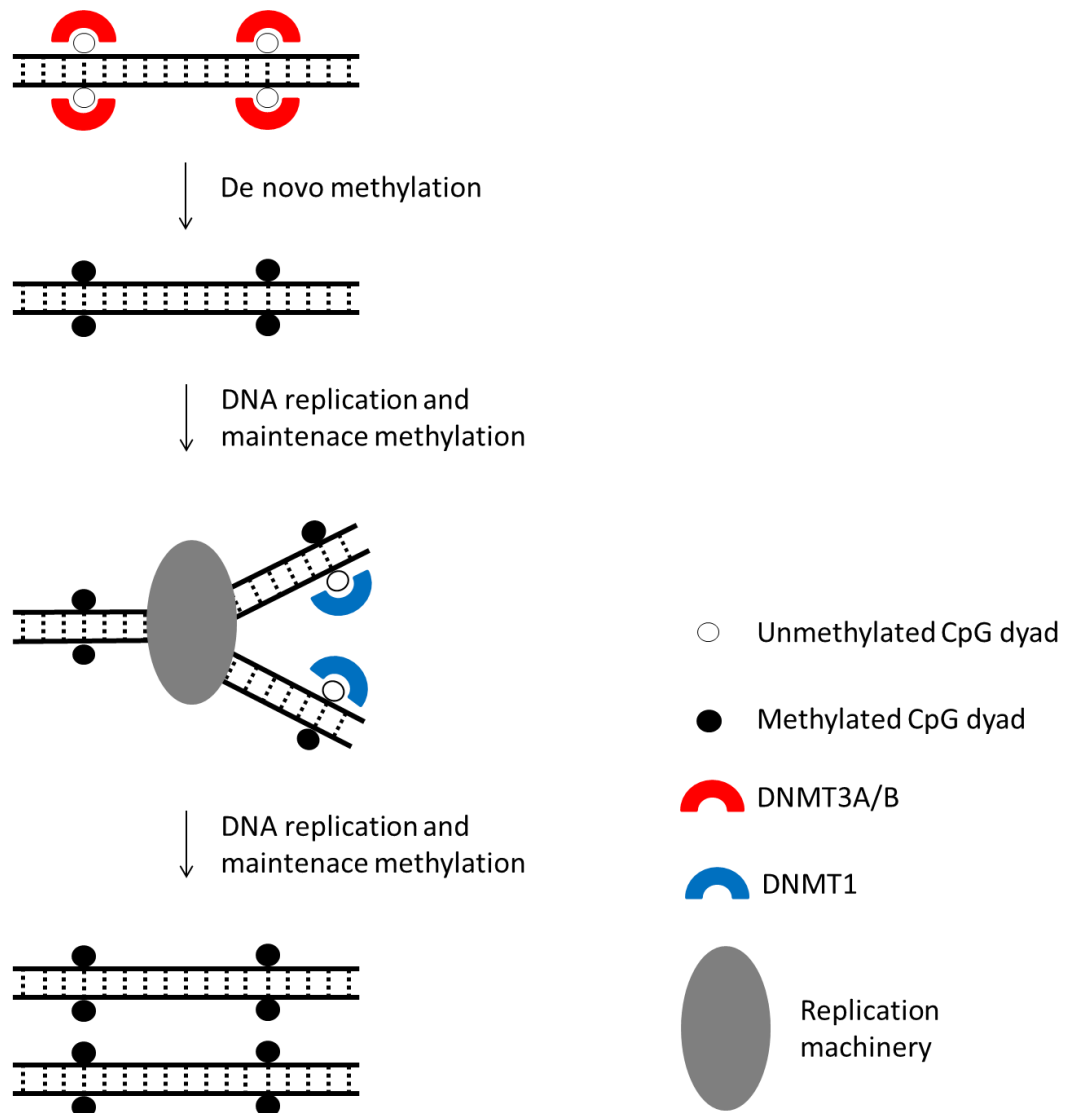


Figure 1.3. Establishment and maintenance of DNA methylation in mammalian cells.

Newly formed methylation patterns are catalysed by DNMT3A/B. Un-methylated, but complementary, single, daughter DNA strand, the result of DNA replication becomes methylated by the maintenance methylase, DNMT1.

The first DNA methyltransferase enzyme, reported in 1988 was DNMT1 (Bestor *et al.*, 1988). This enzyme is considered responsible for maintaining DNA methylation in mammalian cells since it catalyses the reaction in which hemi-methylated DNA became methylated at 5-to-30 fold the rate than that seen for unmethylated DNA (Yoder *et al.*, 1997). DNMT1 was shown to maintain the methylation profile created by DNMT3a and 3b

(Pradhan *et al.*, 1999). During the G1-G2 phase of the cell cycle the enzyme, DNMT1, is found to be diffuse throughout the cell, however, in S phase of the cell cycle it is associated with the replication fork. This observation suggests that DNMT1 plays a role in maintaining DNA methylation pattern through the process of semi-conservative replication and in this way ensures that the hemi-methylated daughter chromosomes are fully methylated (Liu *et al.*, 1998; Leonhardt *et al.*, 1992).

Following the identification and characterisation of DNMT1 the DNA methylases, DNMT3a and DNMT3b were identified as the enzymes principally responsible for *de novo* methylation (Okano *et al.*, 1999). In mouse embryonic stem cells, knock down of these two enzymes led to the blocking of DNA methylation patterns, however, the DNA methylation profiles of the imprinted genes are maintained in these cells (Okano *et al.*, 1999). This data suggests that DNMT3a and DNMT3b are not directly involved in the maintenance of DNA methylation, however, studies suggest that, these two enzyme have overlapping functions in *de novo* and maintenance of DNA methylation during development (Okano *et al.*, 1999).

III. DNA demethylation

Many of the aberrant epigenetic alterations are frequently harmful, however, these changes are reversible. In contrast to the enzyme responsible for methylation those leading to demethylation and, therefore, the removal of the methyl group from 5-mC has not thus far been identified. Indeed, although several DNA methyltransferases have been identified there are still no conclusive reports of *bone fide* DNA demethylating enzyme. Where DNA demethylation does occur it is through a process that leads to dilution through successive cell generations and during the DNA replication process. In these cases, an unmethylated single DNA strand is synthesised and this can result, following a further round of

replication, in both strands of the double helix becoming methylation free (Lim and van Oudenaarden, 2007). However, for this to occur the maintenance DNA methylase (DNMT1) must be inhibited and available evidence suggests that DNMT1 recognises hemi-methylated DNA. Recently, putative mechanism of DNA demethylation by oxidation of 5-mC to 5-hydroxymethylcytosine (5-hmC) has been proposed and briefly reviewed in section 1.1.2.

IV. Abnormal regulation of DNA methylation in cancer

Numerous studies that have focused on the relationship between abnormal epigenetic modifications and cancer have revealed two major phenomena. Global genome wide hypomethylation is a common feature of most cancer types, for review see, (Sharma *et al.*, 2010). Concomitant with this change is the observation of hypermethylation of the normally unmethylated CGIs associated with promoter regions (Figure 1.4). These changes, in promoter associated CGI are common and observed in numerous growth regulating genes and those involved in receptor mediated pathways and apoptosis (Sharma *et al.*, 2010; Portela and Esteller, 2010).

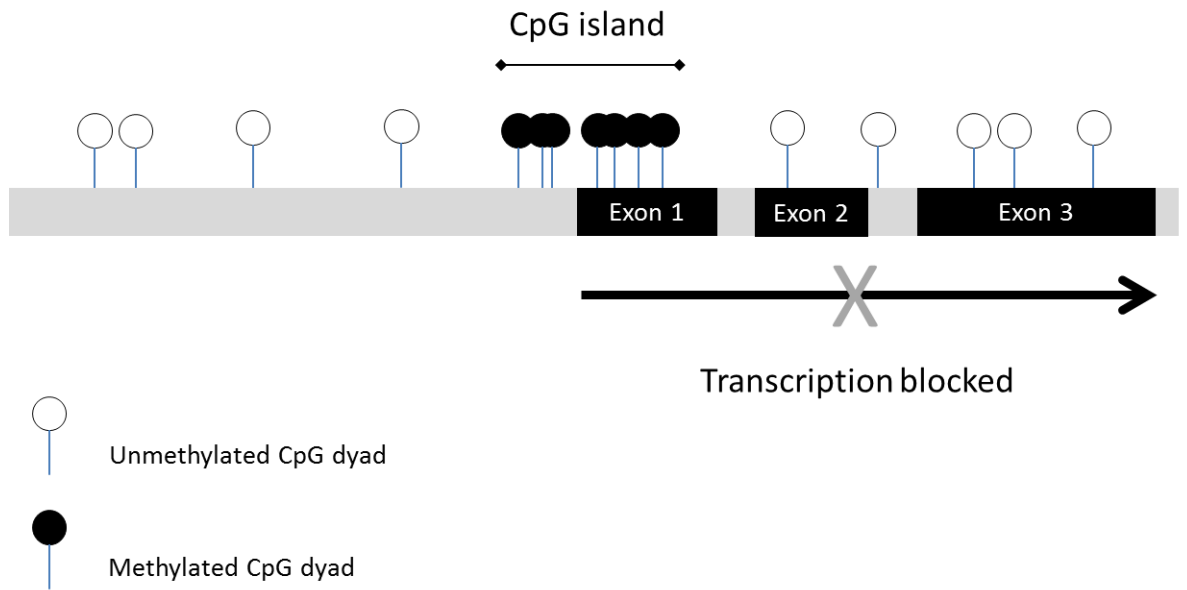


Figure 1.4. Abnormal DNA methylation in cancer.

In contrast to the epigenetic changes (in DNA methylation) apparent in normal cells the DNA methylation profiles in cancer cells are reversed. In these cases, CGIs are methylated but non-CGI CpGs are hypomethylated.

IV.1. Global hypomethylation

Multiple reports have described a link between global, genome-wide hypomethylation of DNA, and a broad spectrum of tumour types. In these cases, most of the genome-wide hypomethylation is found within DNA repeat elements that comprise 45% of the genome (Ehrlich, 2002; Hoffmann and Schulz, 2005). As described previously, this phenomenon is linked to genomic instability and is also associated with reactivation of repetitive DNA elements and oncogenes (Ehrlich, 2002). In cancer cells, data has suggested the link between global hypomethylation, chromatin restructuring and nuclear disorganization precede and are perhaps causal for changes in the activity and expression of the histone-modifying enzymes and chromatin regulators (Hoffmann and Schulz, 2005). Interestingly, although global hypomethylation is a common feature across multiple

tumour types its frequency shows differences between tumour type, tumour stage, and the specific sequences harbouring these changes (Hoffmann and Schulz, 2005). Genome-wide hypomethylation has also been shown, in some, but infrequent, cases, to lead to inappropriate activation of growth-promoting genes such as *SI000A4* in colon cancer (Nakamura and Takenaga, 1998), carbonic anhydrase in renal-cell cancer (Uemura *et al.*, 2000), *H-RAS*, *C-MYS*, *SERPINB5*, *CCND2* in gastric cancer (Fang *et al.*, 1996; Akiyama *et al.*, 2003; Oshimo *et al.*, 2003), and *PAX2* in endometrial cancer (Wu *et al.*, 2005). Moreover, in testicular cancer, hypomethylation has been shown to activate some cancer/testis specific genes (Jang *et al.*, 2001) and the micro-RNA gene *let-7a-3* (Brueckner *et al.*, 2007).

IV.2. Gene specific CpG island hyper- and hypo-methylation

In addition to global hypomethylation, there is an association between gene silencing and gene specific promoter CGIs methylation. Although this association has been reported in a larger number of publications, it is still not entirely clear whether CGIs methylation is a cause or the consequence of gene inactivation (Costello and Plass, 2001). For example, the genes on one of the two X chromosome in females are inactivated, however, the inactivation occurs prior to the associated CGI methylation. Thus, while some studies suggest a causal association between methylation and gene silencing other studies suggest that silencing precedes methylation and that methylation is necessary to maintaining the gene silencing (Gartler *et al.*, 1992).

Similar to global hypomethylation, gene specific CGI methylation is also frequently observed in many cancer types. This phenomenon was first detected in retinoblastoma (Greger *et al.*, 1989; Sakai *et al.*, 1991) as associated with the inactivation of *RBI* that is regarded as a key tumour suppressor genes (TSG) (Ohtani-Fujita *et al.*,

1993). Subsequent to these early studies numerous other TSGs have been identified, where loss of expression is associated with promoter methylation rather than genetic mutation. As example only, these include *CDKN2B* (Herman *et al.*, 1996), the Von Hippel-Lindau gene (*VHL*) (Herman *et al.*, 1994), *CASP8* (Teitz *et al.*, 2000), and *GADD45* (Zerbini and Libermann, 2005) in numerous different tumour types. Epigenetic silencing by methylation of promoter regions are also observed in some genes responsible for DNA repair mechanisms such as *hMLH1*, *MGMT*, and *BRCA1* in sporadic cancers (Esteller, 2000).

Gene silencing, in association with CGI methylation frequently occurs in the early stages of tumour development (Feinberg *et al.*, 2006). Indeed a broad range of genes, including *CDKN2A*, *SFRP1*, *GATA4*, and *GATA5* show increased promoter methylation in tumour associated and histologically normal tissue and in the pre-invasive stages of many cancers whereas genetic mutations are rarely detected at this stage (Issa, 2004).

The majority of gene specific CGIs changes are methylation of previously unmethylated CGI and leads to gene inactivation. However, relatively few publications reported the inverse phenomenon, that is, hypomethylation associated with increased expression of genes with oncogenic properties. As example, the melanoma associated antigen (MAGE) family were found to be overexpression by promoter hypomethylation in testis cancer (De Smet *et al.*, 1996), and MAGE-A3 in pituitary tumour (Zhu *et al.*, 2008a). Another example is the *S100A4* gene encoding a calcium binding protein which is also reactivated in brain tumours by CGI hypomethylation (Singh *et al.*, 2003).

The involvement of hypo- and hyper-methylation mechanisms is summarized in figure 1.5.

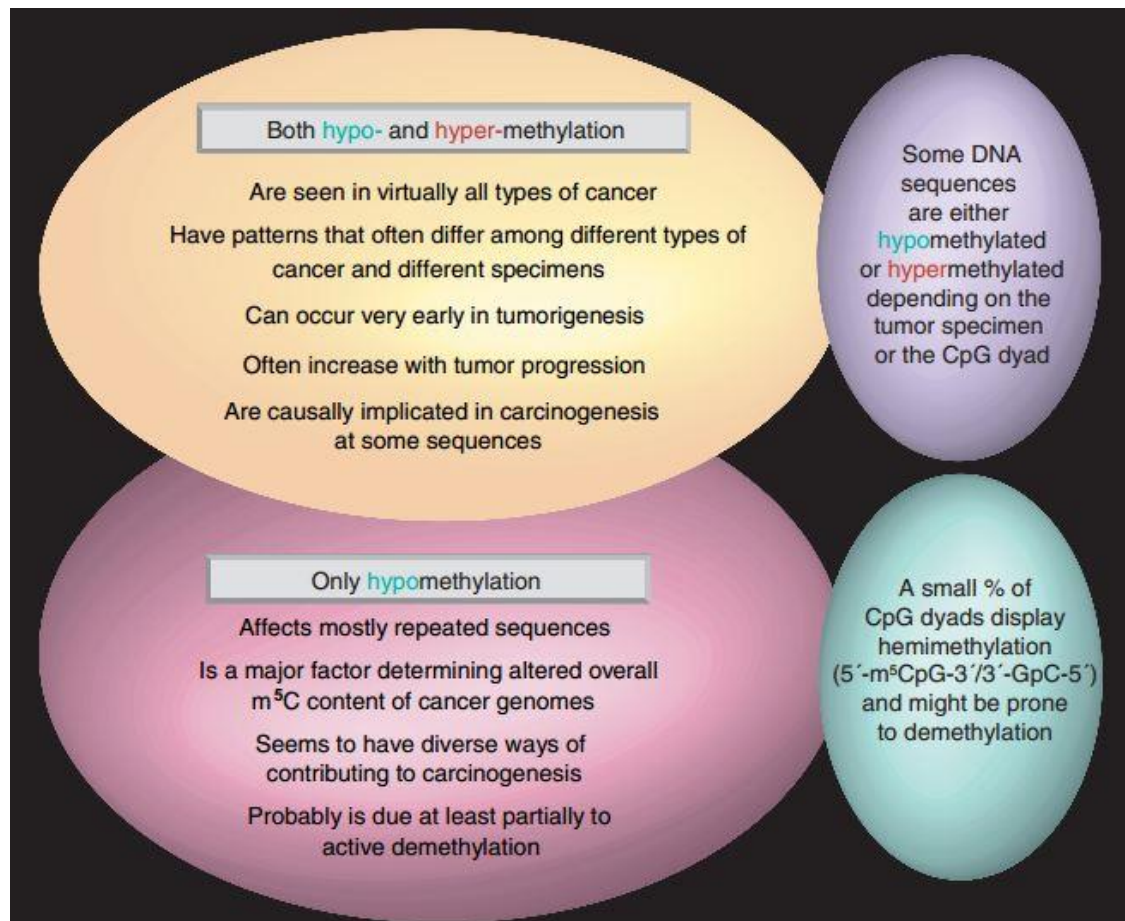


Figure 1.5. Similarities and differences in cancer-associated hypo- and hyper-methylation of DNA.

(Ehrlich, 2009).

IV.3. Aberrant methylation of CpG island shores

It is apparent that DNA methylation of CGIs is frequently associated with promoter regions and this change frequently leads to or reinforces transcriptional silencing (Deaton and Bird, 2011). However and more recently the potential involvement of CpG sites outside, or on the borders of *bone fide* CpG islands has been subject to detailed investigation and has revealed important new understanding. In these studies, the importance of low density CpG dinucleotides, that lie approximately 2 kb up- and/or down-stream of CGIs has been uncovered. In these cases, these regions are named and

referred to as CpG island shores (Irizarry *et al.*, 2009). Similar to CGI, methylation of CpG island shores also leads to gene inactivation and may indeed correlate more closely with these events (Portela and Esteller, 2010). It is perhaps surprising, that the majority of the tissue- and cancer-specific DNA methylation is not specific for the gene associated CpG islands, but rather at CpG island shores (Doi *et al.*, 2009; Irizarry *et al.*, 2009). CpG island shores methylation is conserved between human and mouse and can serve as an epigenetic hallmark to distinguish tissue specificity. Furthermore, the importance of these regions is further reinforced by the findings that, 70% of the differentially methylated regions, implicated in cell type reprogramming, are associated with CpG island shores (Ji *et al.*, 2010; Doi *et al.*, 2009). In addition, for genes with non-CGI associated promoters, these too can be directly silenced by DNA methylation in these shores (Han *et al.*, 2011). These observations suggest that CGI shore methylation may be involved in tissue differentiation, epigenetic reprogramming and cancer. Indeed, and as example, a genome-wide analysis of CpG methylation has shown that CpG island shore, but not CGI methylation, regulates caveolin-1 expression in breast cancer (Rao *et al.*, 2012).

IV.4. Abnormal imprinting

It has been proposed that inappropriate inactivation of the previously active allele at imprinted gene loci can lead to loss of gene expression (Knudson, 1970). In addition to these changes in imprinting status, X-chromosome inactivation, loss of heterozygosity, and DNA mutation of the normally expressed allele can all lead to gene silencing. Thus, and in accord with the “two hit” hypothesis proposed by Knudson, the possibility of inactivation of an imprinted gene is higher than for non-imprinted genes (Garinis *et al.*, 2002). Data has suggested that abnormal DNA methylation of imprinted genes might play a role in the development of cancer. A number of imprinted genes are silenced in cancer cells, such as

PEG1 in breast cancer, *DLK/GLT2* in pheochromocytoma, neuroblastoma, and Wilms' tumours (Astuti *et al.*, 2005).

Silencing of the normally expressed allele of imprinted genes (referred to as loss of imprinting) is often related to the development of various types of cancer (Falls *et al.*, 1999). Moreover, if the imprinted gene is a TSG, loss of imprinting of this gene usually leads to tumourigenesis. For example, 40% of breast and ovarian cancer case found associated with the defects of the non-imprinted allele of a TSG *ARHI* (Yu *et al.*, 1999).

Loss of imprinting in most cases, leads to re-expression of the normally silenced and imprinted allele (referred to as relaxation of imprinting). However, and somewhat counter-intuitively the term loss of imprinting can lead to gene silencing. In this case the normally expressed allele becomes methylated. In those cases where relaxation of imprinting is apparent and for the growth promoting genes and/or oncogenes this change can lead to or promote tumour growth. An example is the autocrine growth factor gene *IGF2* in Wilms' tumours. In normal cells only the paternally derived allele is expressed, however, in tumour cells the paternal and maternal alleles are expressed and this leads to increased *IGF2* expression (Ogawa *et al.*, 1993). In pituitary derived growth-hormone-secreting tumours, relaxation of imprinting of the oncogene *GNAS* is associated with activation of the normally silenced allele (Hayward *et al.*, 2001).

1.1.2. DNA hydroxymethylation

Recently, a further modification of DNA methylation has been described as 5-hydroxymethylcytosine (5-hmC) (Kriaucionis and Heintz, 2009; Tahiliani *et al.*, 2009). Formation of 5-hmC is derived from 5-mC by the activity of chromosome ten-eleven translocation (TET) protein family (Kohli and Zhang, 2013). Oxidization of 5-mC to 5-

hmC is catalysed by TET1 (Tahiliani *et al.*, 2009), TET2 (Ko *et al.*, 2010) and TET3 (Ito *et al.*, 2010). Moreover, TET enzymes can further convert 5-hmC to 5-formylcytosine (5-fC) (Ito *et al.*, 2011) and 5-carboxylcytosine (5-caC) (He *et al.*, 2011; Ito *et al.*, 2011).

Because 5-hmC is found at high frequency in several types of mammalian cells, this epigenetic modification may play important roles in cell functions. For example, in cerebellar Purkinje cell DNA, approximately 20% of all cytosine in CpG dinucleotides are present as 5-hmC (Kriaucionis and Heintz, 2009). Strong evidences are available to support that formation of 5-hmC plays a role as a mechanism of DNA demethylation *in vivo*. Overexpression of TET1 in cultured cells leads to decrease in 5-mC levels together with increase in the unmodified cytosine levels (Tahiliani *et al.*, 2009). The converse phenomenon was observed in mouse embryonic stem cells where knock-down of TET1 leads to the increase in levels of 5-mC (Xu *et al.*, 2011). Furthermore, a two-steps mechanism of replication-independent demethylation has been proposed. The first step is the deamination of 5-mC into thymine by cytosine deaminases of the activation-induced cytidine deaminase (AID)/apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC) family (Zhu, 2009). Alternatively, there is another route in which 5-mC is oxidized to 5-hmC by TET proteins followed by the conversion of 5-hmC to 5-hydroxyuracil (5-hmU). The second step is the excision of the previously formed thymine/5-hmU by DNA glycosylases (thymine DNA glycosylase (TDG), methyl-CpG-binding domain protein 4 (MBD4), and single-strand-selective monofunctional uracil-DNA glycosylase 1 (SMUG1)) followed by base excision repair (BER) which results in unmodified cytosine (Cortellino *et al.*, 2011; Guo *et al.*, 2011).

Data suggested that 5-hmC epigenetic marks may be involved in the development of the human brain, and the abnormal 5-hmC mark may play a role in the molecular

pathogenesis of neurodevelopmental disorders (Wang *et al.*, 2012). In cancer, several pioneer studies have shown that 5-hmC is an epigenetic mechanism. For example, loss of 5-hmC has been shown to be an epigenetic hallmark of melanoma (Lian *et al.*, 2012).

1.1.3. Histone modification

DNA methylation, once established, is a stable modification that silences gene transcription and is inherited through successive cell divisions. However, histone modifications are labile and can be reversed by cellular processes and enzymatic pathways (Sharma *et al.*, 2010). In contrast to DNA methylation, histone modifications can lead to either activation or repression of gene transcription (Hebbes *et al.*, 1988; Kouzarides, 2007). In particular, the N-terminal tails of histones can undergo several different and often concomitant modifications, including methylation, acetylation, ubiquitylation, sumoylation and phosphorylation of key residues (Kouzarides, 2007).

Most descriptions of histone modifications and their influence on gene expression have focused on the changes in acetylation and methylation patterns of lysine residues in histone N-terminal tails (Bilodeau *et al.*, 2006; Ezzat *et al.*, 2006; Al-Azzawi *et al.*, 2011; Jenuwein and Allis, 2001; Hebbes *et al.*, 1988; Kouzarides, 2007). The principal modifications associated with either active or silent genes can be summarised as follows: histone tail acetylation is associated with active genes, as is tri-methylation of lysine 4 in the N-terminal of histone 3 (H3K4me₃); (Liang *et al.*, 2004; Hebbes *et al.*, 1988) however, tri-methylation of lysine 9 and/or lysine 27 in the N-terminal of histone 3 (H3K9me₃ and/or H3K27me₃, respectively) or lysine 20 in the N-terminal of histone 4 (H4K20me₃) are associated with transcriptional repression (Kondo *et al.*, 2008; Berdasco and Esteller, 2010; Fraga and Esteller, 2005) (Figure 1.6). These changes have been termed “the histone code” (Jenuwein and Allis, 2001). Thus, the combinations of changes to the histone tails,

acting either alone or in concert with CpG island methylation, determine gene expression profiles in normal cells. Moreover, the expression status of these epigenetically modified genes frequently affects down-stream signalling pathways and can lead to diseases such as cancer (Egger *et al.*, 2004; Jones and Baylin, 2002). Epigenetic changes, together with genetic alterations, have important roles in tumour initiation, progression and responses to therapeutic interventions (Jones and Baylin, 2002).

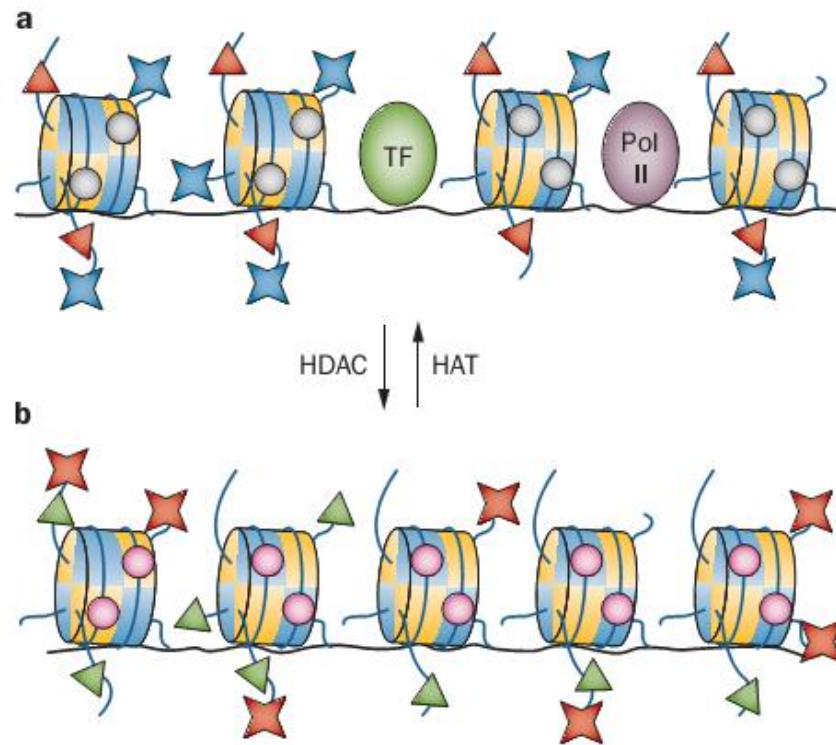


Figure 1.6. Epigenetic modifications associated with transcriptionally active and transcriptionally silent genes.

(a) A transcriptionally active gene with an open, relaxed, chromatin structure. DNA (black line) with non-methylated CpG islands (grey circles) in the promoter regions. The DNA is coiled around the nucleosomes (blue and yellow cylinders). Histone 3 tails (blue lines) are acetylated (red triangles) at lysine 9 (H3K9Ac) or trimethylated (blue stars) at lysine 4 (H3K4me3). The relaxed chromatin configuration allows access by transcription factors and associated transcriptional machinery, including RNA polymerase II.

(b) A silent gene with a closed chromatin structure. The promoter CpG islands are methylated (red circles), and the histone 3 tails are trimethylated (red stars) at lysine 9 (H3K9me3) or lysine 27 (H3K27me3). Accompanying these changes in methylation, the histone tails are hypoacetylated (green triangles) and these patterns are associated with closed chromatin.

Abbreviations: HAT, histone acetyl transferase; HDAC, histone deacetylase; Pol II, RNA polymerase II; TF, transcription factor. (Yacqub-Usman *et al.*, 2012b)

Multiple members of the histone methyl transferases (HMTs) family have been described and one of these, EZH2, is overexpressed in breast and prostate tumours and leads to aberrant H3K27me3 (Valk-Lingbeek *et al.*, 2004). Acting in concert with, and opposing the actions of the HMTs, is a family of lysine specific demethylases that, together with the HMTs, maintain global histone methylation patterns. Inappropriate expression of lysine specific demethylases has been reported for some tumours (Metzger *et al.*, 2005; Shi *et al.*, 2004; Cloos *et al.*, 2008). In one such example, LSD1, a lysine demethylase efficiently removes both activating, histone 3 lysine 4 methylation (H3K4me) and repressive, histone 3 lysine 9 methylation marks (H3K9me), depending on its specific binding partner (Metzger *et al.*, 2005; Shi *et al.*, 2004). In this case, therefore, the binding partner dictates the role of LSD1 as either a co-repressor or co-activator of gene expression.

1.1.4. Mechanistic interdependence between DNA methylation and histone modifications

The histone changes associated with condensed chromatin and gene silencing are frequently accompanied by, or associated with, DNA methylation (although not invariably). The mechanism by which DNA methylation might directly regulate gene expression is not entirely clear. Indeed, chromatin remodelling, as described above, is principally mediated through repressive histone modifications (Figure 1.6). However, the observation that transcriptional silencing is frequently associated with gene promoter methylation might reflect the interdependence of these two processes (Figure 1.7). Examples of how these processes are integrated are apparent from multiple studies showing that methylated DNA sequences preferentially bind a family of methyl-binding domain proteins that recruit HDACs to these sites (Figure 1.7) (Wade, 2001). Similarly,

DNMT1, which maintains DNA methylation patterns, also binds to HDACs (Fuks *et al.*, 2000; Rountree *et al.*, 2000; Robertson *et al.*, 2000). Conversely, at nonmethylated DNA sequences, the associated histones are acetylated by histone acetylases (HATs) that contribute to an open chromatin structure conducive to gene transcription (Belinsky *et al.*, 2003; Jones and Baylin, 2002).

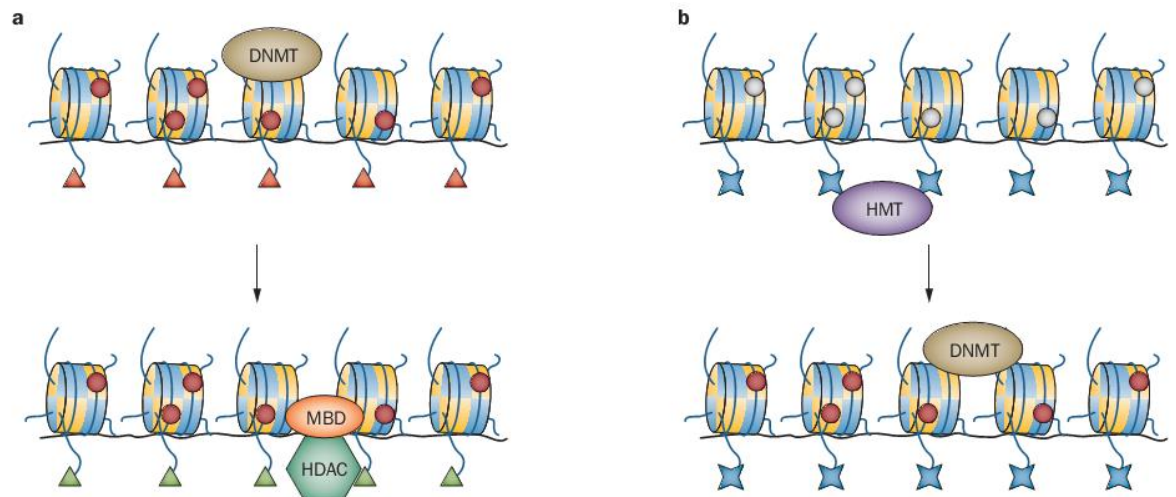


Figure 1.7. Interdependence of DNA methylation and histone modifications.

(a) Interaction between DNA methylation and histone deacetylation. DNA CpG island methylation (red circles) by DNMTs precedes binding of MBDs, which recruit HDACs to induce transcriptional repression by mediating deacetylation (green triangles) of histone tails and chromatin remodelling.

(b) Interaction between histone 3 lysine methylation and DNA methylation. Nonmethylated CpG islands in promoters (grey circles) are accompanied by histone 3 trimethylation at lysine 4 and lysine 27 (H3K4me3 and H3K27me3; blue stars) mediated by the HMTs precedes de novo DNA methylation (red circles) by DNMTs.

Abbreviations: DNMT, DNA methyltransferase; HDAC, histone deacetylase; HMT, histone methyltransferase; MBD, methyl-CpG binding domain protein. (Yacub-Usman *et al.*, 2012b).

Although the interdependence of DNA methylation and histone modification has not been directly investigated in pituitary cells and tumours, the siRNA-mediated knockdown of DNMT3B transcription in pituitary cells provides some insight into this

relationship (Zhu *et al.*, 2008b). DNMT3B is thought to be a *de novo* DNA methylase, but its knock-down in pituitary cells is associated with increased histone acetylation and decreased histone methylation (Zhu *et al.*, 2008b). No reports indicate any direct effect of this enzyme on histone acetylation or methylation, which suggests that changes in DNA methylation status in this experimental model indirectly affect histone modifications. Studies of the promoter CpG islands associated with *D2DR* and *BMP-4*, and of histone modifications in these regions also suggest that these changes are indeed integrated (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a).

The opposing actions of HATs and HDACs on covalent histone modifications contribute to the reversible nature of acetylation and deacetylation patterns and to their subsequent effects on gene expression (Figure 1.7). However, other histone tail modifications, such as methylation, also show widespread changes in both normal cells and tumour cells (Sharma *et al.*, 2010). Methylation of histone tails is mediated by histone methyl transferases (HMTs). Although the principal function of these enzymes is histone modification, they are also able to recruit the DNA-methylating enzyme DNMT1 to these regions (Figure 1.7). In this way, and in contrast to DNA-directed histone modifications, histone methylation can drive CpG island methylation.

1.1.5. Pharmacological reversal of epigenetically silenced gene

The epigenetic aberrations that distinguish tumour cells from their “normal” counterparts are, in essence, reversible (Yang *et al.*, 2010; Kelly *et al.*, 2010). Indeed, a broad range of drugs, the “epidrugs” that target the epigenome and effectively reverse DNA methylation and histone modifications have been described (Karagiannis and El-Osta, 2006a; Gore *et al.*, 2006; Cortez and Jones, 2008). Many of these drugs have as their target the enzymes responsible for these aberrations, that is the DNMT’s and HDACs

(Sharma *et al.*, 2010). Their principle mode of action is to either lead to the “trapping” of enzymes (DNMTs) on to the DNA, leading to their relative depletion within cells or, for the HDACs, to directly inhibit enzyme activity. For some of these drugs US Food Drug Administration (FDA) approval has resulted in their use for the treatment of, as example, myelodysplastic syndrome (Plimack *et al.*, 2007; Sharma *et al.*, 2010).

I. DNMT inhibitors

The two approved drugs that target the DNMTs are nucleoside analogues, and marketed as Azacidine and Decitabine, while still others are in various stages of clinical trials. For these analogues there are some concerns regarding toxicity, particularly as high doses (Egger *et al.*, 2004). However, since these drugs should target mainly dividing tumour cells they are thought to exert minimal effect on other slowly dividing cells (Yang *et al.*, 2003). The most significant issues and concerns associated with the use of these particular drugs is that they are they are chemically unstable in aqueous solutions. Secondly, they have also been shown to suppress the proliferation of blood cells from the myeloid lineage and thereby lead to toxicity problems (Kantarjian *et al.*, 2003). However, other nucleoside analogues, including zebularine (see below) are much more stable in aqueous solution and this drug is also reported to preferentially target tumour cells rather than their non-malignant counterpart (Cheng *et al.*, 2004b; Cheng *et al.*, 2003). Other drugs in development that reverse DNA methylation achieve their inhibitory effects by directly blocking enzyme activity (Mai and Altucci, 2009).

II. HDAC inhibitors

Similar to the spectrum of drugs that target methylation, a range of useful and promising drugs have been described that target and inhibit the HDACs. Most of these

drugs exert their effect by blocking the enzymes catalytic domain and thereby blocking substrate recognition and binding. It is beyond the scope of this review to discuss the various HDAC inhibitors and the reader is directed toward an excellent review of this area by Kristensen and colleagues (Kristensen *et al.*, 2009). However, and of note, one of these inhibitors, trichostatin A (TSA), that displays considerable efficacy as an inhibitor of deacetylation (Yoshida *et al.*, 1990), has not progressed to clinical trial because of its severe side effects. Although several other HDAC inhibitors are at various phases of clinical trial (Mai and Altucci, 2009) the only inhibitor with US FDA approval is, vorinostat, also known as SAHA, for the treatment of cutaneous T-cell lymphoma. A frequent concern is that the HDAC inhibitors might exert global effects on gene expression; however, this does not appear to be the case since several microarray studies have shown them to affect only a small fraction of the transcriptome (Chiba *et al.*, 2004; Dannenberg and Edenberg, 2006).

III. Combinations of DNMT and HDAC inhibitors

The histone changes associated with condensed chromatin and gene silencing are frequently, but not invariantly, accompanied by, or associated with, DNA methylation. Indeed, as previously discussed, the latter can direct the former and vice versa (Figure 1.7). This frequent mutual interdependence has led to exploration of combinatorial drug treatments. Cameron and co-workers first showed a synergistic effect through combined and sequential challenge with a DNMT and HDAC inhibitor (Cameron *et al.*, 1999). Subsequent studies by other groups have confirmed and extended these findings (Klisovic *et al.*, 2003; Belinsky *et al.*, 2003). In fact our own studies that are discussed in a subsequent section also used combined drug challenges. In this case, in a pituitary context the demethylating agent zebularine and a HDAC inhibitor, TSA, to induce expression of epigenetically silenced genes (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a).

IV. Combinations with conventional therapies

Several studies, principally of haematological malignancies treated with DNA methylation inhibitors, report significant improvement of overall survival compared with more conventional treatment regimens. The efficacy of epigenetic therapies as part of a combined therapeutic approach with chemotherapeutic agents has recently been reviewed (Gore and Hermes-DeSantis, 2008; Fenaux *et al.*, 2009). An extension to these types of studies is suggested by Kim and colleagues where approaches using DNMT inhibitors and/or HDAC inhibitors before chemotherapy are reported to be a promising therapeutic option (Kim *et al.*, 2003). In addition, other combined strategies, in this case where epigenetic therapy preceding radiotherapy are also reported and have been shown to increasing the radiosensitivity of tumour cells (Karagiannis and El-Osta, 2006a; Karagiannis and El-Osta, 2006b). To date there are relatively few reports that describe the effects of epigenetic therapy in combination strategies with chemo- or radio- therapy in solid tumours.

V. Genome-wide and gene-specific reactivation of silenced genes

V.1. Genome wide studies

Several groups have exploited epidrugs to reverse epigenetic modifications for identification of genes subject to epigenetic regulation. In general, epidrug challenges, either alone or in combination lead to expression of previously silenced genes. In these types of investigation, genome-wide expression is commonly characterised by cDNA microarray interrogation and has/is used to identify hitherto unknown, and in some cases, tumour-specific genes. Interestingly, and perhaps surprisingly, there appears to be little overlap of genes when different de-methylation and or re-acetylation agents are compared

(Gius *et al.*, 2004; Flotho *et al.*, 2009). In addition, some genes that are re-expressed following drug challenges are not methylated prior to challenge and still others show reduced expression (Gius *et al.*, 2004; Flotho *et al.*, 2009). These findings may reflect epigenetic silencing of an upstream regulator that may in turn stimulate or repress its downstream and “array-identified” targets (Zhu *et al.*, 2001; Soengas *et al.*, 2001; Komashko and Farnham, 2010). Our laboratories investigations in this arena initially employed AtT20 cells challenged with the demethylating agent, 5-aza-2'-deoxycytidine (5-Aza-2-Dc), with the intention of interrogating gene expression by cDNA microarray. However, we found that cell cytotoxicity, at least in AtT20 cells, represented a significant problem and experiments were discontinued (unpublished data). Instead, we initially used a siRNA knock-down approach targeting the maintenance methylase, DNMT1, to “reveal” silenced genes (Revill *et al.*, 2009; Dudley *et al.*, 2008). These types of approaches, either drug or genetic based respectively hold significant promise for identification of novel genes and potential drug targets in multiple tumour types including those of pituitary origin.

V.2. Pharmacologic unmasking for gene specific investigations

Given that epimutations in most tumour types including pituitary adenomas can be functionally equivalent to genetic loss or inactivating mutation, our laboratory has investigated this phenomenon as a potential mechanism for silencing of the dopamine receptor (D2R) and of the cytokine bone morphogenic protein, BMP-4 (Yacqub-Usman *et al.*, 2012a). Our rationale for studying these genes was that some pituitary tumours show reduced expression of these genes and their protein products (An *et al.*, 2003; Giacomini *et al.*, 2006). Indeed, in other tumour types, as example, breast and prostate cancer, loss or reduced expression of the alpha and beta isoforms of the estrogen receptor, ER α / ER β

respectively is associated with CpG island methylation (Li *et al.*, 2000; Lapidus *et al.*, 1996). In these cases, treatment of tumour cells with inhibitors of DNA methylation, azacitidine, and/or in combination with TSA resulted in re-expression of estrogen receptor, restoration of receptor activity and re-sensitization to antiestrogen therapy (Yang *et al.*, 2000; Yang *et al.*, 2001; Walton *et al.*, 2008).

As a first line of investigation and as proof of principle in a pituitary context and we used the rodent pituitary cell line, GH3, that does not expressed the endogenous D2R. However, given our previous experience with 5-Aza-2-Dc and cytotoxicity (see above) we employed zebularine, reported by others to be a less toxic demethylating agent (Cheng *et al.*, 2003; Cheng *et al.*, 2004b). Our lab's studies showed that zebularine in combined challenges with TSA is synergistic in the induction of D2R in GH3 cells. Drug challenges were associated with or were responsible for partial demethylation of the D2Rs CpG island and led to histone modification associated with active genes (Al-Azzawi *et al.*, 2011). Studies by other groups have also shown that partial, zebularine induced, demethylation is sufficient to induce gene re-expression (Cheng *et al.*, 2003; Cheng *et al.*, 2004a; Scott *et al.*, 2007). Similar studies of the BMP-4 promoter also showed that loss of BMP-4 expression in pituitary cell lines and in primary tumours was also associated with epigenetic changes (CGI methylation and histone modifications) (Yacqub-Usman *et al.*, 2012a). As in our laboratories studies of the D2R promoter we found that combined "epidrug" challenges restored gene expression and led to relaxation of chromatin as determined by change in histone modification profiles (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a).

1.2. The pituitary gland

In the vertebrate brain, the pituitary, or master gland, is an endocrine organ located at the base of the hypothalamus. The pituitary gland and the hypothalamus are connected to each other by the median eminence through the pituitary stalk. The pituitary gland is about the same size as a pea and weighs approximately 0.5 g. There are two lobes, the anterior and posterior lobe. Both of the lobes are regulated by the hypothalamus. The anterior lobe comprises at least five hormone secreting cell types including somatotrophs, gonadotrophs, lactotrophs, thyrotrophs, and corticotrophs (Dworakowska and Grossman, 2009). Hormones secreted by the pituitary gland play important roles in the regulation of many physiological processes such as metabolism, growth, sexual development, and response to stress (Dworakowska and Grossman, 2009).

1.2.1. Pituitary adenomas

Pituitary adenomas account for approximately 10–25% of all intracranial neoplasms and are the third most common intracranial tumours in surgical practice (Scheithauer *et al.*, 2006). Pituitary adenomas have been observed approximately at the ratio 1/1000 in the human population (Daly *et al.*, 2007). There are no difference of the tumour incidence between sexes, and similar to most other tumour types it increases with age. Most pituitary adenomas are benign although approximately one third will show invasive growth characteristics. Pituitary carcinomas represent only 0.2% of these adenomas (Scheithauer *et al.*, 2006).

The different types of pituitary adenomas are shown in table 1.1. The most frequent type of pituitary adenomas are prolactinomas (PRL-omas) which represent approximately 50% of cases. Non-functioning pituitary adenomas (NFPAs) – accounts for about 30%,

somatotroph adenomas (GH-omas) 15–20% and corticotroph adenomas (ACTH-omas) account for 5–10% and thyrotroph adenomas (TSH-omas) less than 1% (Lindholm *et al.*, 2001; Levy, 2008). Gonadotrophic adenomas are rare.

Once developed, pituitary adenomas can invade upwards into brain and/or downwards into the cavernous sinuses resulted in the compressive effect. Symptoms associated with this invasion including headaches, visual defects, and cranial nerve dysfunction. Invasion of pituitary adenomas into the brain usually leads to death. While invasion causes similar effects among pituitary adenomas patients, symptoms associated with cell subtype-specific effects are caused by or associated with over-secretion of a particular hormone (Table 1.1) and are varied.

Type of adenoma	Secretion	Pathology
Lactotrophic adenomas	Prolactin	Galactorrhea, Hypogonadism, Amenorrhea, Infertility, and Impotence
Non-functioning adenomas	None	Mainly mass effects and can result in hypopituitarism
Somatotrophic adenomas	Growth hormone (GH)	Acromegaly
Corticotrophic adenomas	Adrenocorticotrophic hormone (ACTH) and Pro-opiomelanocortin (POMC)	Cushing's disease
Thyrotrophic adenomas	Thyroid-stimulating hormone (TSH)	Hyperthyroidism
Gonadotrophic adenomas	Luteinizing hormone (LH) and Follicle-stimulating hormone (FSH)	None

Table 1.1. Types of pituitary adenomas.

1.2.2. Epigenetic changes in pituitary tumour cells

Tumours of pituitary origin display a diverse array of epimutations that affect the transcription of hormone and growth factor receptors, members of the signal transduction pathway, transcription factors and cell-cycle regulators. These epigenetic changes usually (but not invariably) include changes in the DNA methylation status of promoter-associated CpG islands and/or modifications to the underlying histones.

Several key genes are aberrantly expressed in sporadic pituitary adenomas, and the epimutations responsible for their aberrant expression have been identified. Table 1.2 details the major genes subject to epigenetic change in this tumour types. In the majority of cases the genes described, with limited exceptions, have been identified by candidate gene approaches. In these cases investigations have focused on genes previously shown, in other tumour types, to impact upon tumourigenesis. In addition, early studies of the pituitary tumour epigenome and indeed in other tumour types, focused primarily on gene silencing associated with DNA methylation (Simpson *et al.*, 2002; Simpson *et al.*, 2000; Simpson *et al.*, 1999; Woloschak *et al.*, 1997; Farrell, 2006). However, subsequent studies have included characterisations of histone modifications (Al-Azzawi *et al.*, 2011; Yacub-Usman *et al.*, 2012a; Zhu *et al.*, 2007a). These changes are frequently interdependent, and rarely affect gene expression as independent modifications.

I. Cell-cycle regulators

As summarised in Table 1.2, key tumour suppressors, including *CDKN2A* gene, which encodes the p16 protein, is silenced in the majority of pituitary-derived adenomas, and silencing is associated with CGI methylation (Simpson *et al.*, 1999; Woloschak *et al.*, 1997). Furthermore, enforced expression of p16 in pituitary cells reveals that this protein

as a *bona fide* tumour suppressor (Frost *et al.*, 1999). Furthermore, a proportion of growth-hormone-secreting pituitary adenomas, which show reduced expression of retinoblastoma protein have shown this to be associated with CGI methylation or discrete sequence deletions within the protein-binding pocket (Simpson *et al.*, 2000). Other genes with reduced or complete loss of expression, and associated with increase CGI methylation, include *MEG3* (Zhao *et al.*, 2005), *GADD45γ* (Zhang *et al.*, 2002), and *NNAT* (Dudley *et al.*, 2008; Revill *et al.*, 2009). Enforced or induced expression of these genes in pituitary and other tumour cell lines inhibits cell proliferation. Moreover, additional roles for some of these genes have been described; for example, the *MEG3* transcript that is not translated induces p53 expression (Zhou *et al.*, 2007). Interestingly, *MEG3* and *NNAT* are genomically imprinted in normal pituitary cells and most other tissues.

II. Signal transduction pathways

In pituitary tumours, abnormal expression of genes related to signal transduction pathways have been reported. As example, the guanine nucleotide-binding protein Gsα subunit is a component of numerous signal transduction pathways and is encoded by the *GNAS* gene (also known as GSP). In this case, inappropriate expression is confined to growth-hormone-secreting adenomas. Activation is associated with genetic mutation at one of two discrete sites in this gene. However, subsequent to the description of activation, through genetic mutation, inappropriate expression of Gsα has also been shown to be associated with relaxation of imprinting at the *GNAS* locus (Hayward *et al.*, 2001). Moreover, relaxation of imprinting is only observed in tumours that do not harbour the activating mutation, suggesting that these aberrations are mutually exclusive. Pro-apoptotic factors, including the so named, pituitary tumour apoptosis gene (*PTAG*), is silenced by CGI methylation at high frequency across the majority of pituitary adenoma subtypes

(Bahar *et al.*, 2004b). Enforced expression studies in pituitary tumour cells augments bromocriptine mediated apoptosis. The *DAPK1* gene is also silenced through CpG methylation in pituitary tumours, although this epimutation seems to be confined to the invasive adenoma subtype (Simpson *et al.*, 2002).

III. Hormone and growth factor receptors

Table 1.2 also describes hormone and growth factor receptors where inappropriate expression is associated with epimutation. These include the fibroblast growth factor receptor 2 (*FGFR2*) gene transcript that is subject to alternative splicing and reduced expression of this isoform (FGFR2-IIIb) is a result of DNA and histone methylation (Abbass *et al.*, 1997). Re-expression of FGFR2 can be induced in pituitary cells following exposure to a DNA demethylating agent and a histone deacetylase inhibitor (Zhu *et al.*, 2007b). Of note, in tumour cells of both pituitary and thyroid origin, a reciprocal expression pattern exists with the normally silenced *MAGEA3* gene, which encodes melanoma-associated antigen A3 (MAGE-A3) (Zhu *et al.*, 2008a; Kondo *et al.*, 2007). Interestingly, inappropriate expression of *MAGEA3* is associated with DNA hypomethylation, and both FGFR2-IIIb and MAGE-A3 rely on DNA hypo-methylation and histone modification to inter-regulate each other's expression (Asa and Ezzat, 2002; Zhu *et al.*, 2008a).

Inappropriate increases in expression of the *de novo* DNA methylase enzyme DNMT3B have also been reported in pituitary tumours. Interestingly, increased expression is associated with changes in histone modifications rather than DNA methylation status (Zhu *et al.*, 2008b). Knockdown of DNMT3B (Zhu *et al.*, 2008a) decreases cell proliferation and induces expression of other cell cycle regulators including p21, p27 RB1

These findings suggest that DNMT3B might exert broad epigenetic control on the transcriptome.

IV. Transcriptional regulators

A dominant-negative isoform of the transcription factor Ikaros, Ik6 is expressed in a proportion of pituitary adenomas (Ezzat *et al.*, 2006). This protein increases both acetylation of histone H3 and expression of the anti-apoptotic gene *BCL-XL*

Gene*	Mode of activation or inactivation	Effect on expression	Product	Function
<i>GNAS</i>	Mutation or relaxation of imprinting	Increase	Guanine nucleotide-binding protein Gs α subunit (short isoform)	Signalling pathways, oncogene
<i>MAGEA3</i>	Hypomethylation or relaxation of imprinting	Increase	Melanoma-associated antigen 3	Cell cycle regulation, oncogene
<i>DNMT3B</i>	Histone modification	Increase	DNA (cytosine-5)-methyltransferase 3B	De novo methylase
<i>CDKN2A</i>	CpG island methylation	Silencing or reduction	Cyclin-dependent kinase inhibitor 2A, isoforms 1/2/3 (p16INK4a)	Cell cycle regulation, tumour suppressor
<i>RB1</i>	CpG island methylation or deletion	Silencing or reduction	Retinoblastoma protein	Cell cycle regulation, tumour suppressor
<i>DAPK1</i>	CpG island methylation or deletion	Silencing or reduction	Death-associated protein kinase 1	Apoptosis
<i>MEG3</i>	Promoter or enhancer methylation	Silencing or reduction	Very putative protein from MEG3locus ‡	Cell cycle regulation, tumour suppressor
<i>GADD45G</i>	CpG island methylation or deletion	Silencing or reduction	Growth arrest and DNA damage-inducible protein GADD45 γ	Cell cycle regulation, tumour suppressor
<i>RHBDD3</i> (also known as <i>PTAG</i>)	CpG island methylation or deletion	Silencing or reduction	Rhomboid domain-containing protein 3 (also known as pituitary tumour apoptosis gene)	Apoptosis

<i>NNAT</i>	CpG island methylation or deletion	Silencing or reduction	Neuronatin	Cell cycle regulation, tumour suppressor
<i>FGFR2</i>	DNA and histone methylation	Silencing or reduction	Fibroblast growth factor receptor 2 (IIIb isoform)	Cell surface receptor; inhibits proliferation and induces apoptosis
<i>D2R</i>	CpG island methylation and histone modification	Silencing or reduction	Dopamine D2 receptor	Cell surface receptor; induces apoptosis
<i>BMP4</i>	CpG island methylation and histone modification	Silencing or reduction	Bone morphogenetic protein 4	Bifunctional, cell-type-specific stimulates/inhibits proliferation

Table 1.2. Epigenetic changes associated with pituitary tumours.

(adapted from Yacqub-Usman et al., 2012b)

*This list is not exhaustive, and does not include genes with inappropriate expression for which the associated mechanism is not known.

‡No protein has been found. This gene probably only encodes an RNA.

1.3. Identification of novel genes involved in pituitary tumourigenesis

As described in previous sections, epimutations play key roles in the development of most tumour types including those of pituitary origin. It is, therefore, important to develop efficient high-throughput techniques to explore epimutations in tumour cells. To date, numerous studies in this and other tumour type, that have focused on epigenetic change have been published, however, earlier studies were confined or restricted to identification on a gene-by-gene (single gene approach) that examined perhaps one or a limited number of candidate genes. Therefore, genome-wide epigenome profiling techniques were a natural successor for these types of study.

1.3.1. Limitation and advances in studies of epigenetic changes associated with pituitary adenomas

I. Candidate gene approach

For epimutations that characterise pituitary adenomas, their identification, in early studies has been through use of candidate gene approaches (Dudley *et al.*, 2009). Less frequently, for identification of novel genes, investigators have used differential display techniques. These approaches have employed either tumour-derived cDNA (Zhang *et al.*, 2003) or DNA that has been subject to prior methylation-sensitive digestion (Bahar *et al.*, 2004b). Although these techniques have identified novel genes, a limitation is that identification is frequently confined to a single transcript or gene, that is, on a gene-by-gene basis. With the exception of several repetitive elements such as LINE-1, a further limitation of single gene approaches is that they are unable to provide a picture of genome-

wide epigenetic changes such as region specific hypomethylation and subtype-specific methylation patterns.

II. Subtype specific

In common with specific genetic aberrations which frequently show subtype specificity, epimutations also show a degree of specificity (Simpson *et al.*, 2000; Zhang *et al.*, 2002; Zhao *et al.*, 2005). As an example, subtype-specific genetic changes are apparent for the *gsp* oncogene and are also apparent for inappropriate expression of HMGA2 in somatotroph- and lactotroph-derived adenomas, respectively (Farrell and Clayton, 2000; Hayward *et al.*, 2001; Fedele *et al.*, 2006).

Genome-wide studies employed differential display technique (Zhang *et al.*, 2002; Zhang *et al.*, 2003) have identified novel genes involved in pituitary tumourigenesis. Although this technique was not focused on epigenetic change, the mechanism of GADD45 γ silencing was subsequently shown to be through DNA methylation (Bahar *et al.*, 2004a). A further limitation of this technique is that it can only compare the cDNA library of a subtype specific adenoma (monoclonal cell population) with the cDNA of normal pituitaries (polyclonal cell populations). As described in section 1.1.2, of the total cells of a normal pituitary, the number of subtype-specific cells is limited. Therefore, the novel genes identified by comparing the two unequal cDNA libraries may not accurately reflect or represent their expression levels. If this is the case, then epigenetic modifications might not impact on expression. To overcome this problem our lab has conducted a siRNA knock-down study of *DNMT1* in the pituitary cell line AtT20 cell line (Dudley *et al.*, 2008) and compared findings to these cells treated with a non-target siRNA. In this cases, while comparing cell that were like-for-like they were restricted to the corticotroph cell types.

III. Species boundaries

Of the few genome-wide studies in pituitary tumourigenesis using siRNA knock-down or pharmacological unmasking strategies (Al-Azzawi *et al.*, 2011; Dudley *et al.*, 2008), a limitation of these techniques is that they are reliant upon actively dividing cells for the effective reversal and/or erasure of epimutations. Moreover, primary pituitary adenomas show limited proliferative potential *in vitro* and investigations are, therefore, reliant upon pituitary adenoma cell lines of mouse or rat origin. While these approaches have uncovered novel genes, they are constrained by the issue of extrapolation across species boundaries, and their promoter methylation status had to be assessed by further experiments (Dudley *et al.*, 2008).

IV. Limitation of clinical sample

As described in the section 1.1, the limitation and constrain of tumour size, where most tumours measure no more than 5mm in diameter is formidable and particularly for genome wide investigations. Since epigenetic studies frequently require significant amounts of samples for characterisation of multiple changes (modifications) and include: histone modifications, DNA methylations, and expression gene transcripts and protein expression techniques for whole genome amplification are necessary. A subsequent section of this thesis will describe the optimisation and validation of this technique in this tumour types.

1.3.2. DNA methylation profiling strategies

Considering the factors and limitations described above, genome-wide DNA methylation profiling provides significant advantage for identification of novel genes involved in pituitary tumourigenesis. The rapid evolution of genome-wide techniques in recent years has provided researchers with the ability to perform genome-wide DNA methylation profiling that can identify multiple novel genes involved in tumourigenesis. . Second, the limitation of monoclonal versus multi-clonal cell populations which affect mRNA differential display technique can be overcome. In this case, and by default, CGI in normal pituitary is in general not methylated while in tumours, assuming it is associated with gene silencing, will be methylated, however, this association is not invariant. Thus, comparing CGI methylation levels between a particular tumour subtype (monoclonal cell population) with the normal pituitary (multi-clonal cell population) will reflect the actual epigenetic changes that are implicated in or casual in the evolution or progression of a tumour. Moreover, by looking directly at DNA methylation, this reduces the number of other experimental manipulations required to identify mechanism of gene silencing. Third, the challenges imposed by extrapolating across species boundaries is also surmounted and that beset pharmacological unmasking strategy. Finally the limitation imposed by restricted amounts of clinic specimens can be partly overcome by employing whole genome amplification technique.

At the start of my thesis work, and as already eluded to a number of genome-wide DNA methylation profiling techniques were available. To select the appropriate approach, it was first necessary to review the principles of each of the techniques and to determine their relative advantages and disadvantages. The major techniques in the context of my studies are described and discussed below.

1. Underlying methodologies for detection of DNA methylation

1.1. Sodium bisulphite conversion

Classic dideoxy DNA sequencing is not able to detect DNA methylation either directly and particularly since this information is erased by PCR. Therefore, as a first step of DNA sequencing based on, or able to interpret DNA methylation methods for converting DNA or more specifically, DNA bases that are not methylated whilst leaving those that are methylated resistant to conversion. In this case the technique termed sodium bisulphite sequencing achieves this outcome through converting unmethylated cytosine to uracil which will subsequently become thymine after PCR amplification (Clark *et al.*, 1994). In contrast, the methylated cytosine is not converted. Therefore, in the final sequencing reaction the unmethylated and methylated cytosines are distinguishable as T and C, respectively. Since the initial report in 1992 (Frommer *et al.*, 1992), bisulphite conversion has become the strategy for analysis of DNA methylation.

A remarkable advantage of bisulphite conversion is that it then allows use of several technique, including direct dideoxy sequencing and array based analysis, for the detection at the single base level. However, a concern of bisulphite-conversion-based methods is the incomplete conversion rate which may have critical effects in the accuracy and reliability of the results (Grunau *et al.*, 2001; Genereux *et al.*, 2008). However, most techniques now result in success rates of between 99.5% to 99.7% (Grunau *et al.*, 2001).

1.2. Methylation sensitive restriction enzymes

The majority of the methylation-sensitive endonucleases are inhibited by methylation of their recognition sites. For example, *HpaII* recognise the sequence CCGG, and its endonuclease activity is blocked by methylation of either cytosine (Korch and

Hagblom, 1986). However, there are also a very limited number of enzymes for example *McrBC*, that specifically digest methylated sequence (Sutherland *et al.*, 1992). Therefore, digestion of DNA with a methylation-sensitive restriction enzyme is another underlying methodology to detect DNA methylation. This methodology enables the possibilities of either discriminating and/or enriching for particular methylated or unmethylated fractions of DNA throughout the genome (Khulan *et al.*, 2006; Lippman *et al.*, 2005). Because the recognition sequence of each specific enzyme is known, an advantage of this technique is resolution at a single base level. However, compared to sodium bisulphite which allows detection of methylation at any CpG, detection possibility of methylation-sensitive enzymes is limited to CpG lying within their specific recognition sequences. In addition, and similar to sodium bisulphite methodology, digestion reaction by endonucleases can also face the problem of incomplete digestion that lead to biased methylation result. This issue is especially important in those cases where a PCR step is involved in the downstream methylation profiling procedure. Furthermore, and of particular importance in the area of pituitary research, methods based on methylation-sensitive endonuclease digestion usually require high quality, quantity, and integrity of input DNA (Laird, 2010).

1.3. Methylated DNA immunoprecipitation

The third underlying methodology for DNA methylation analysis is the affinity purification of methylated or unmethylated DNA fragments using antibodies that recognise specific modifications. Commercially available antibodies that recognise methylated cytosines can be used to enrich methylated DNA sequences (Weber *et al.*, 2005; Gebhard *et al.*, 2006). An alternative to antibody enrichment is by using recombinant protein containing methylated DNA binding domains (MBD) (Rauch *et al.*, 2006; Gebhard *et al.*, 2006). Methylated DNA fragments can also be purified by an indirect approach that

employs chromatin immunoprecipitation technique using antibodies against MBDs (Ballestar *et al.*, 2003). The major limitation of affinity-based methods is relatively low resolution since they can only look at methylation status of a DNA region instead of single CpG site resolution.

II. Genome-wide DNA methylation profiling strategies

Following use of one or several of the pre-treatment methods described above various types of genome-wide DNA methylation analysis strategies can be used. A summary of the major DNA methylation profiling strategies are shown in table 1.3. Furthermore, an appraisal of sample throughput versus genome coverage of each of the major techniques has been subjected to recent review (Laird, 2010) and is shown in figure 1.8. In general, the evolution of global DNA methylation profiling can be classified into three stages (Laird, 2010). The first stage employed classical methods, including gel electrophoresis and Sanger dideoxy sequencing, to read out DNA methylation information. The second stage uses array techniques. The third stage takes the advantages of the Next Generation Sequencing (NGS) technologies. It is noteworthy that the classification is not always applicable since combinations of the methods are also used in some cases.

Pretreatment method	Resolution	Advantages	Disadvantages
Bisulphite conversion	High	<ul style="list-style-type: none"> • Applicable to any samples 	<ul style="list-style-type: none"> • Require complete conversion • Degradation of DNA
Methylation-sensitive restriction enzyme	Moderate	<ul style="list-style-type: none"> • Highly sensitive • Simple 	<ul style="list-style-type: none"> • Analysis limited to methylation at restriction sites • Complete digestion required • Require high quality, quantity, and integrity of input DNA
Affinity purification	Moderate	<ul style="list-style-type: none"> • Good for studies focus on CGIs and repetitive sequences 	<ul style="list-style-type: none"> • High 5mC density required • Only suitable for rapid, large scale, low resolution DNA methylation studies

Table 1.3.1. Genome-wide DNA methylation analysis: sample pretreatments

Readout method	Theoretical genome coverage	Sample throughput	Advantages	Disadvantages
Oligo-nucleotide Arrays	Genome-Wide	Low to moderate	<ul style="list-style-type: none"> Provides a “snap shot” of methylation within specific Tiled regions, such as promoters Custom Arrays options available Bioinformatic and analytical requirements are less challenging relative to NGS 	<ul style="list-style-type: none"> Require <i>a priori</i> knowledge of the regions to be probed Cross-hybridisation between similar sequences
BeadArray (Illumina)	~450,000 CpG sites	Moderate	<ul style="list-style-type: none"> Compatible with low levels of input DNA No PCR bias Good snap shot of methylation at key promoter- associated CpG site and other regions Reproducible results Bioinformatic and analytical requirements are less challenging relative to NGS 	<ul style="list-style-type: none"> Requires <i>a prior</i> knowledge of the regions to be probed
Dideoxy sequencing following sodium bisulphite conversion	Genome-Wide	Low	<ul style="list-style-type: none"> Technically simple Single base resolution 	<ul style="list-style-type: none"> Large number of “hands-on” experiments Usually requires T:A cloning and multiple sequencing runs Time consuming High cost and labour intensive for genome-wide study
Next generation sequencing (NGS)	Genome-Wide	Low	<ul style="list-style-type: none"> Fast Genome-wide single base resolution 	<ul style="list-style-type: none"> High cost Bioinformatics technically demanding

Table 1.3.2. Genome-wide DNA methylation analysis: readout methods

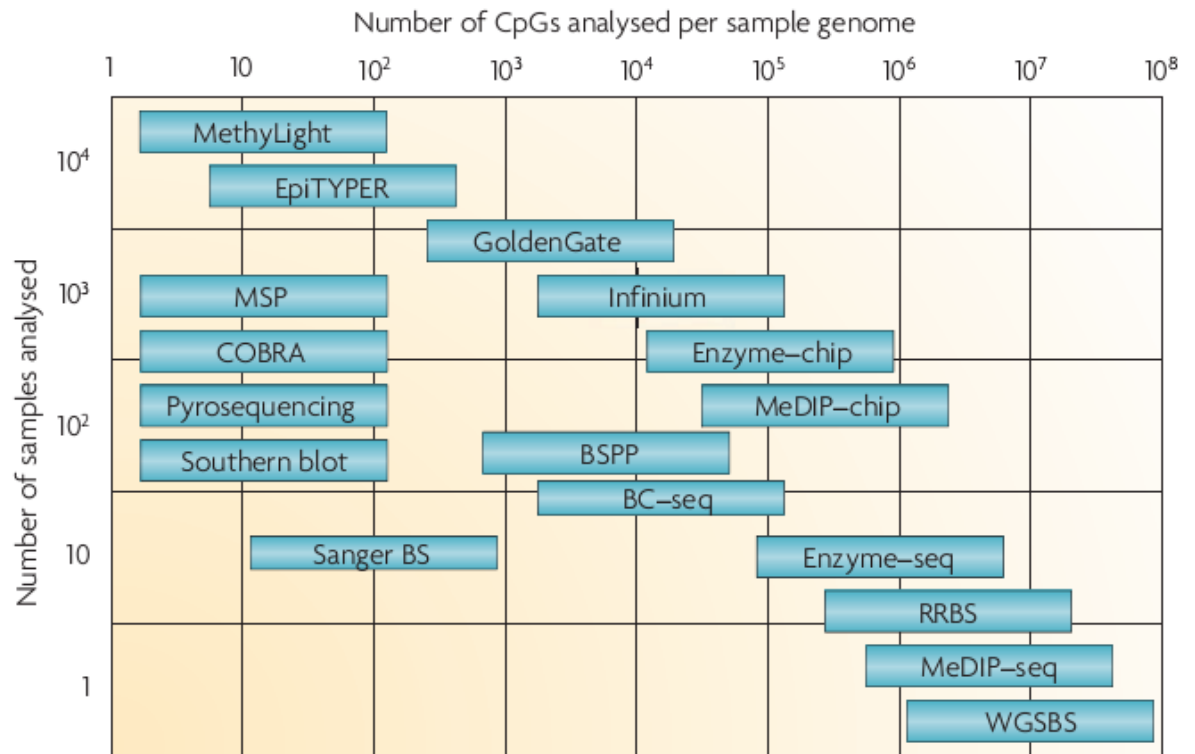


Figure 1.8. Sample throughput versus genome coverage.

(Laird, 2010)

A plot of sample throughput against genome coverage for various DNA methylation techniques. Throughput is determined by the number of samples that can be analysed per experiment, based on large eukaryotic genomes. Coverage is determined by the number of CpGs in the genome that can be analysed per experiment. BC-seq, bisulphite conversion followed by capture and sequencing; BS, bisulphite sequencing; BSPP, bisulphite padlock probes; -chip, followed by microarray; COBRA, combined bisulphite restriction analysis; MeDIP, methylated DNA immunoprecipitation; MSP, methylation-specific PCR; RRBS, reduced representation bisulphite sequencing; -seq, followed by sequencing; WGSBS, whole-genome shotgun bisulphite sequencing

II.1. Classical strategies

Gel-based strategies

The first genome-wide DNA methylation profiling study was reported in 1991 which developed a gel-based strategy named restriction landmark genomic scanning (RLGS) (Hatada *et al.*, 1991). Briefly, genomic DNA was treated with methylation-sensitive restriction enzymes which preferably digest certain regions such as CGIs, followed by two dimensional gel electrophoresis. Differences in methylation are reflected by different patterns of restriction fragments visualized on the gel. To date, RLGS has been used in ~100 studies according to Pubmed search for “DNA methylation restriction landmark genome scanning”. The major disadvantage of this strategy is labour intensive and requires further steps to identify differently methylated fragments. In addition, sample requirement of RLGS is ~2 µg of high-purity, high-molecular-weight genomic DNA. Compare to most other techniques, this sample size requirement is higher than most other technique or technologies (reviewed by Laird, 2010). For example, Illumina Infinium 27K assay requires less than 1 µg genomic DNA to profile >27,000 CpG sites, while Illumina GoldenGate is reported to be compatible with degraded DNA extracted from formalin-fixed, paraffin-embedded samples (Killian *et al.*, 2009). To date, the maximum number of CpG islands interrogated in a single RLGS experiment is reported at ~1000 (Costello *et al.*, 2000) which is relatively small compare to modern array- or NGS-based strategies.

Another gel-based strategy is Methylation-sensitive Arbitrary Primed PCR (MsAP-PCR) (Liang *et al.*, 2002). Similar to RLGS, this technique also employs methylation-sensitive enzyme to digest genomic DNA. Digested fragments were amplified by arbitrary primers following by visualization by gel electrophoresis to compare the differences between the samples. Our lab used this technique coupled with differential display to

compare the pool of normal pituitaries with tumours subtypes. As result, we found a novel gene, *PTAG*, aberrantly methylated in non-functioning tumours but not in normal pituitaries. Subsequent functional studies reported its involvement in pituitary tumourigenesis *PTAG* (Bahar *et al.*, 2004b). Although this technique allows isolation of novel aberrantly methylated/unmethylated DNA fragments, the following cloning and sequencing steps to identify these fragments is gene-by-gene basis.

Bisulphite sequencing based strategies

In the absence of high throughput NGS techniques, classical approaches using Sanger dideoxy sequencing of bisulphite converted DNA to profile methylation status of target regions. The major problems of the Sanger sequencing technique for analysis of methylation status is that it requires a cloning step, usually a T:A vector, and each sample requires 10 to 20 clones to be subject to individual sequencing, thus significantly increasing experimental manipulations and costs. This issue is highly critical in genome-wide study. To overcome these issues, an integrated pipeline which allows direct high-throughput bisulphite sequencing of PCR products (Rakyan *et al.*, 2004) has been developed that allows for the elimination of the cloning steps. In addition to this approach, a bioinformatics algorithm has been applied to calculate methylation values of the directly sequence PCR fragments, and also to adjust for incomplete conversion signals (Lewin *et al.*, 2004). This method was employed to interrogate human chromosomes number 6, 20, and 23 in 12 different tissues (Eckhardt *et al.*, 2006). The most comprehensive database of methylation generated by pre-NGS bisulphite sequencing strategies is the Human Epigenome Project. To date, methylation values of ~2 million CpG sites have been published online at www.epigenome.org. Despite these developments, genome-wide studies using classical Sanger bisulphite sequencing technique still require significant “hands-on” experimental time and high cost consumable budgets.

Considering the size of human genome, another issue of whole-genome bisulphite sequencing is that a large proportion of sequenced regions do not contain CpG dyad. Therefore, a method that allows genome-wide analysis to focus on CpG rich regions is desirable (Suzuki and Bird, 2008). Reduced-Representative Bisulphite Sequencing (RRBS), a restriction-enzyme-based method, has been developed for this purpose (Meissner *et al.*, 2005). Briefly, genome DNA is first digested by a restriction enzyme which recognizes a CpG-containing sequence. This ensures that every clone to be sequenced contain at least one CpG site. The following steps include size-selection and isolation of digested fragments by gel electrophoresis, ligation to a linker, sodium bisulphite conversion, PCR amplification, and cloning to generate a reduced representative library of the genome. This library is then subsequently sequenced. Early studies have necessitated thousands of sequencing reactions (Meissner *et al.*, 2005) which are not always applicable to smaller laboratories or where cost or labour are a limiting factor.

II.2. Array hybridization based strategies

Technological evolution apparent through the advent and use of microarray based techniques has enabled the interrogation of a larger number of CpG sites in a single experiment. Basically, there are two major kinds of arrays based on their probe types including region specific array and tiling array. Region specific array provide direct answers with respect to the methylation status of specific regions, however, these techniques are not suitable to interrogation of DNA methylation at a genome-wide level. In contrast, tiling arrays can in principal interrogate methylation across the genome (in a region or tile specific manner), but require multiple arrays to cover large genome such as that found in *homo sapiens*. For these reasons and for the technical constraints and limitation NGS techniques provide significant advantage(s). Indeed, from a tiling array

perspective these also require multiple further steps to identify aberrantly methylated regions. To date, several types of array assays are commercially available. In considering the number of probes offered by these commercial arrays, as example >450,000 CpG sites on the Illumina Infinium Methylation Assay 450K, these region specific arrays are preferred. However, it is important to emphasise that although these would be regarded as region specific this specificity relates primarily to the absolute number of CpGs analysed across the genome.

The first array-based genome-wide DNA methylation study was conducted in 1999 (Huang *et al.*, 1999). Differential methylation hybridization (DMH) method had in this case had been developed to profile methylation of CpG islands in breast cancer. Briefly, CpG island probes are fixed on the glass slide. Genomic DNA of normal and cancer samples are firstly treated with *MseI* which cut the genome frequently but rarely in CGI, and secondly with the methylation-sensitive endonuclease *BstU1*. Digested fragments are ligated with linkers and subsequently amplified by PCR using primers targeted to the linker sequences. As a result, different methylation patterns between samples are now reflected as different amplified amplicons. These amplicons are labelled and then applied onto the array chip to read out the differences of methylation. This technique was employed to profile the pool of 28 primary breast tumours paired with corresponding normal tissues. Number of probes reported in this study is 1104 (Yan *et al.*, 2000). Several modifications/developments based on principle of DMH have been reported. Instead of methylation-sensitive enzyme, the methylation-dependent enzyme *McrBC* has been used for efficient profiling of densely methylated regions (Pietrobono *et al.*, 2002).

It is noteworthy that DMH can only answer “yes – no” question but is not quantitative. This issue can be addressed by array-based methods utilize affinity enriched samples. In the technique named Methylated DNA Immunoprecipitation (MeDIP) (Weber

et al., 2005), genomic DNA is randomly sheared to fragments, denatured, and enriched by a monoclonal antibody against 5-mC. As a result, only methylated DNA fragments are enriched. The non-enriched (input) and enriched samples are differently fluorescence labelled and subsequently applied on to the chip for competitive hybridization with probes. Fluorescence density between input and enriched sample are calculated and methylation levels of each specific target sequence can be quantitatively read out. To enhance the signal strength and/or deal with limited source of DNA sample, enriched and input fragments can be subjected to a whole genome amplification (WGA) step prior to labelling (Ren *et al.*, 2000). There are several alternatives of MeDIP such as methylcytosine immunoprecipitation (mDIP or mCIP) and methylated CpG island recovery assay (MIRA).

Illumina Infinium Methylation assay

At the time when my studies began a high throughput, quantitative approach for genome-wide profiling of DNA methylation had become available for researchers as the so named Illumina Infinium Methylation assay. The initial array based technology marketed by this company interrogate 1536 CpG sites per sample and 96 samples per run was known as the GoldenGate (Bibikova *et al.*, 2006). Subsequently, the so named, 27K array were made available that interrogated 27,578 CpG sites across 14,495 protein-coding gene promoters and 110 microRNA gene promoters and allow for the simultaneous interrogation across 12 samples on a single “chip” (Bibikova *et al.*, 2009). This array includes almost 13,000 genes in the NCBI Consensus CDS Database (Genome Build 36). Most of the important loci of nearly 1,000 cancer-related genes reported in literature have been covered. Highest attention, by increasing the number of CpG per gene, was paid to 144 established cancer genes which showed abnormal DNA methylation status. Within each of the CGI, optimal CpG sites were chosen by i) holistically analysing the anticipated

assay performance, ii) location in CGI, and distance to the transcriptional start site. Data has proven that in one CGI the methylation status of one specific CpG site often correctly reflects the methylation status of other closely surrounding CpG sites (Eckhardt *et al.*, 2006). Therefore, it seems sufficient to evaluate the methylation status of a CGI by accessing its two representative CpGs. It was this array that has been used in my study. However, more recently a 450K array has become available but this was not commercially available until 2011.

The technique itself comprises two specific “bead” types: The first bead type (labelled U) comprise probes designed to match, that is base pair, with unmethylated site; the second bead type (labelled M) containing probe designed to match with a methylated site (Figure 1.9). To distinguish between these two possibilities, that is a methylated or unmethylated DNA, the genomic DNA is first subjected to sodium bisulphite conversion. In this case, and after conversion C residues (not in the context of a guanine, that is not as a CpG) are then converted to T. This conversion ensures that the “probe” is specific for converted DNA. However, for C residues in the context of a CpG their conversion is dependent upon their methylation status. As described earlier, cytosine that are not methylated will be converted whereas, cytosine that are methylated will be resistant to conversion and remain as C residues (Figure 1.10). Each of the bead-associated probes is then subject to single base pair extension. As shown in figure 1.10, it is this extension with a differentially fluorescently labelled single base that generates the signal. The relative values are then computationally calculated and converted to Beta value to indicate the methylation state of desired locus (Bibikova *et al.*, 2009).

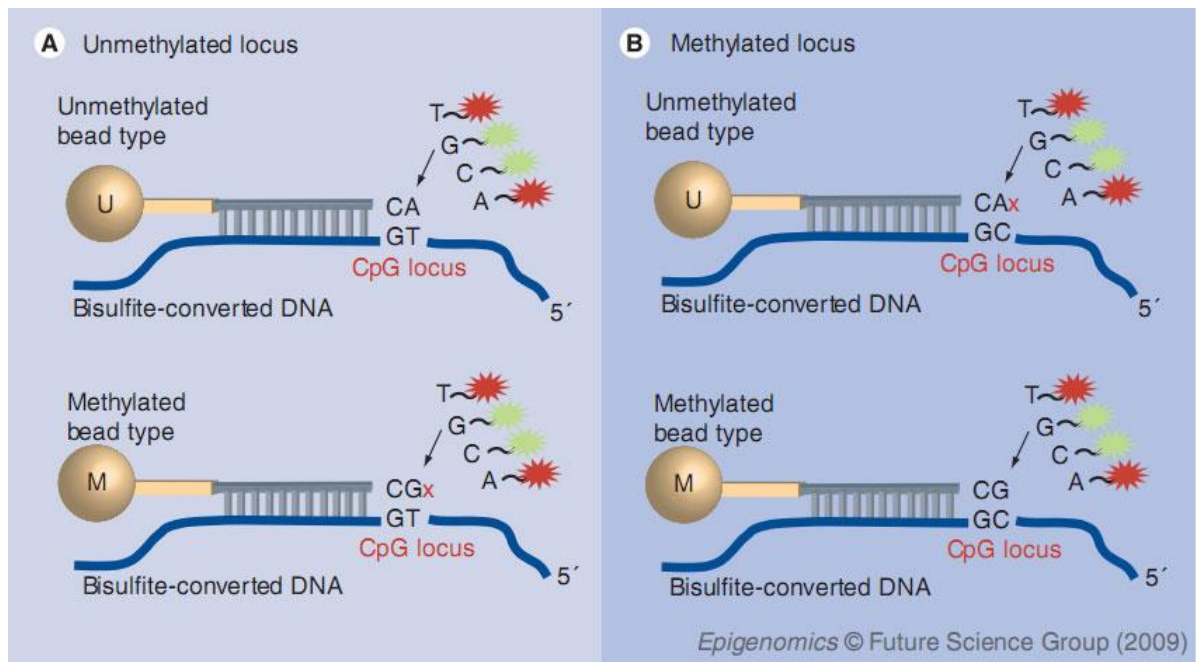


Figure 1.9. Principle of Illumina Infinium Methylation assay.

(Bibikova *et al.*, 2009).

Since the technology does not require a PCR amplification step it avoids selective bias towards short fragments. It looks at an average of 2 CpG sites, each in discrete sites within a CGI to provide genome-wide coverage of methylation patterns. Interestingly, the data generated by Infinium can be integrated with other platforms such as gene expression and microRNA profiling. Moreover, and of high importance to pituitary research, sample requirement of 27K Beadarray is $\sim 1 \mu\text{g}$ of genomic DNA is smaller than most of other genome-wide DNA methylation profiling techniques (Laird, 2010).

II.3 NGS-based strategies

There are ~ 55 million CpG dyads in each human diploid cell (Martin-Subero and Esteller, 2011). Therefore, although array-based platforms have significantly been revolutionized the ability to profile genome-wide DNA methylation they are not applicable

or suitable to generate information at single CpG resolution. The largest array-based platform, Illumina 450K, interrogates ~450,000 CpGs which represents less than 0.1% of CpGs across the genome. However, the development of NGS platforms now enables us to interrogate CpGs at the single base level across the genome. To date, there are several commercially available high-throughput genome wide sequencing platforms. These platforms include the Roche Genome Sequencer FLX, Illumina Genome Analyzer, ABI Sequencing by Oligonucleotide Ligation and Detection. Pre sequence sample treatments and methodologies are broadly similar as to those for arrays, and with the incorporation of minor modifications can then be used for NGS platforms. To distinguish array-based and sequencing-based strategies derived from the similar sample treatment process, the suffix – seq is added. For example, MeDIP-seq is the NGS approach of MeDIP.

As described in section 1.3.2.II.1, the major issues associated with classical Sanger sequencing for analysis of DNA methylation is that cloning and sequencing of multiple clones are required. This limitation has largely been circumvented by the pyrosequencing technique that was initially developed by Mostafa Ronaghi (Ronaghi *et al.*, 1996; Ronaghi *et al.*, 1998). The application of this technology and its application to epigenetic studies is usually termed bisulphite pyrosequencing (Colella *et al.*, 2003). This technique is ideal for DNA methylation analysis at not only the single-gene level but also applicable to genome-wide investigations (coupled with genome-wide platforms). Pyrosequencing significantly reduces man power and cost compared to classical Sanger sequencing and the core methodology of pyrosequencing is based on a sequencing-by-synthesis technology in which nucleotide incorporation is measured using a luminescence signal generated by pyrophosphate release. Bisulphite-converted DNA is first amplified by PCR where one of the two primers is biotinylated. The subsequent PCR products is then “captured” and immobilized on streptavidin beads, then denatured to ssDNA and sequenced. A significant

limitation of pyrosequencing is read lengths where, and at best, sequence reads are restricted to 300-500 nucleotides which is at least half the length of Sanger dideoxy sequencing. This issue is exacerbated in bisulphite pyrosequencing where read lengths are restricted to 25–30 bases. This limitation has been partially improved by modifications of the protocol to allow sequence reads of up to 75 bp (Dupont *et al.*, 2004). Based on a similar principle to that of bisulphite pyrosequencing, the Roche GS FLX has been developed (Margulies *et al.*, 2005). Each single DNA molecule is amplified by emulsion PCR, following by attached to beads and then placed in a microtitre plate for subsequent pyrosequencing. The fluorescence released in the reactions is calculated and converted to methylation values. To date, it is able to process 400,000 reactions of approximate 250 bp and incorporate data from 100 Mb of sequence information per run. Roche GS FLX is currently one of the strongest platforms for profiling cytosine methylation by shotgun sequencing of bisulphite-treated DNA.

The major challenge of genome-wide DNA methylation profiling by NGS techniques is the reduced sequence complexity after bisulphite treatment. In this case, the genetic code that comprised four bases is reduced in complexity to three bases since most of cytosines, except for methylated ones, are converted to thymidines. This challenge has been partly overcome by employing RRBS and/or different capture methods of specific DNA sequence before or after bisulphite conversion (Ball *et al.*, 2009; Hodges *et al.*, 2009; Li *et al.*, 2009). Also, current NGS methods have not yet been able to interrogate the recently discovered 5-hmC modifications. A protocol has been developed to profile 5-hmC (Song *et al.*, 2012), and undoubtedly this will attract interest to be adapted to NGS platforms.

The current state of the art with respect to genome-wide DNA methylation profiling is by whole-genome shotgun bisulphite sequencing (WGSBS) for *Arabidopsis thaliana*

(Cokus *et al.*, 2008) and *Homo sapiens* (Lister *et al.*, 2009; Laurent *et al.*, 2010). However, the high cost and limited number of sample per run of NGS (Figure 1.8) should be considered for each study purpose. Different aspects of available genome-wide DNA methylation strategies are overviewed in table 1.3. In addition, because of the huge amount of raw data which is measured in gigabyte or even terabyte unit, processing of this data usually needs considerable computational infrastructure and professional bioinformatics and/or statistics human power.

1.3.3. DNA hydroxymethylation profiling strategies

It is challenging to detect and quantify 5-hmC because the majority of methods used for studying 5-mC do not distinguish this modification from 5-hmC. In the bisulphite reaction 5-hmC is not converted to uracil but to cytosine 5-methylenesulfonate (CMS). Therefore, 5-hmC is undetectable by bisulphite sequencing because it is remained as cytosine in the final sequencing read (Huang *et al.*, 2010). However, the combination of bisulphite conversion and immunoprecipitation of CMS allows detection of this modification form (Pastor *et al.*, 2011). Moreover, DNA fragments with high density of CMS shows low efficiency in PCR amplification because of the stalling of taq DNA polymerase (Huang *et al.*, 2010). In some circumstances, this leads to underestimation of 5-hmC level with bisulphite-based strategies. Difficulty was also observed with methods based on methylation sensitive restriction enzymes where the activity of most of enzymes is inhibited by both 5-mC and 5-hmC presented in their recognition sites (Tahiliani *et al.*, 2009). In case of affinity-based methods, the constrain is that the affinity is specific to 5-mC but not to 5-hmC (Jin *et al.*, 2010).

I. Glucosylation methods

Although there are constraints in detecting 5-hmC, several strategies of detecting and quantifying 5-hmC have been developed. The first strategy uses bacteriophage T4 beta-glucosyltransferase (BGT) to glucosylate 5-hmC. This method, in combination with MspI restriction enzyme which digests C^mCGG and C^{hm}CGG, but not C^{gluc}CGG recognition sites, has enabled locus-specific 5-hmC quantification by quantitative PCR (Ficz *et al.*, 2011). Moreover, BGT-based strategy also allows genome-wide quantification of 5-hmC by adding a radiolabeled substrate in the glycosylation reaction and measuring the amount of this indicator relative to standards (Szwagierczak *et al.*, 2010). Another affinity-based method named GLIB is the combination of glucosylation, periodate oxidation, and biotinylation to add biotin to 5-hmC. Subsequently, this complex is purified by streptavidin and sequenced by either Helicos or Illumina (Pastor *et al.*, 2011; Song *et al.*, 2011a). Sensitivity of GLIB was reported at 90% of enriched DNA fragments containing at least one 5-hmC (Pastor *et al.*, 2011).

II. Affinity enrichment methods

Affinity enrichment methods are divided into two groups according to the corresponding antibodies which can detect either 5-hmC or CMS. The efficiency of 5-hmC antibodies is dependent to the density of 5-hmC. This property causes inefficiency of the method in some tissues where the level of 5-hmC is low (Pastor *et al.*, 2011). Based on the principle of MeDIP-seq, hMeDIP-seq has been developed (Ficz *et al.*, 2011; Xu *et al.*, 2011). Compared to anti-5-hmC methods, the anti-CMS antibodies are less density dependent and more sensitive, therefore CMS enrichment results in lower background

noise. CMS enrichment couples with Illumina NGS platform has been reported (Pastor *et al.*, 2011).

1.4. Aims of this study

Although there are specific and important exceptions, genetic aberration associated with or perhaps responsible for the outgrowth of primary sporadic human pituitary adenomas are an infrequent finding. However, in marked contrast to the paucity of genetic aberration, and at the outset of my research, it is apparent that a propensity toward epigenetic changes is a feature and inherent property of this tumour types. At the instigation of this project several studies had shown that these changes are manifest as inappropriate methylation of CpG islands (Dudley *et al.*, 2009; Yacqub-Usman *et al.*, 2012b). However, more recently, in this and indeed other tumour types, epigenetic change apparent as histone modification, either alone or in combination with CpG island methylation has been reported (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a; Yacqub-Usman *et al.*, 2012b).

For the majority of genes identified, as showing epigenetic change in this tumour types, a candidate gene approach had been adopted. This approach, and other less frequent approaches designed to identify novel genes are subject to extensive review in the Introduction to this thesis. Therefore, in this thesis the principle questions I wished to address related to characterisation of genome-wide changes in methylation. However and in this case, although essentially genome-wide, it had as its focus CpG islands and in particular those associated with gene associated promoter regions. Furthermore an aim was to perform this analysis in each of major pituitary adenoma subtypes and determine the effect or consequences of these changes on gene expression. The longer term aims were to identify specific epigenetic signature that characterised each of the adenoma subtype and perhaps identify common genes responsible for tumour outgrowth or those particular to a

subtype. Finally a further aim, assuming novel genes were identified, was to determine their functions through use of cell line model systems.

To perform these studies it was necessary to use what was at that time a relatively new technology, namely Illumina Infinium 27K Array in combination with sodium bisulphite converted DNA from each adenoma subtype. However, it was first necessary to overcome the technical limitation imposed by limited tumour material for the analyses. To circumvent this issue it was necessary to use whole genome amplification (WGA). However, since this technique was relatively untried it was first necessary to determine not only the relative fold amplification but also the correlation between analysis prior to and following WGA. In a similar manner it was also necessary to validate data derived through BeadArray analysis and conventional bisulphite sequencing. To address these aims receptor pathways associated with dopaminergic regulation (D2 receptor [D2R]) and the cytokine (bone morphogenic protein 4 [BMP-4]) pathway were investigated in pituitary cell lines and primary tumours. The epigenetic aberrations examined included CpG island methylation and chromatin remodelling, in this case manifest as histone tail modifications.

Section B: Materials and Methods

Chapter 2: Material and methods

2.1. Primary Tissue material

2.1.1. Normal Pituitary Tissue

Pathological normal mouse pituitaries (CD-1) and rat pituitaries (Sprague Dawley) were obtained from a commercial source, Charles River Research laboratories (Kent, UK).

As control four normal human pituitaries were used. These were post-mortem normal pituitaries (NP) acquired within 12 hours of death with no evidence of endocrine disease. The normal pituitaries were pulverised under liquid nitrogen using a Biopulveriser stainless steel mortar and pestle device (Biospec, California, USA). The resulting granular admix was then stored as 30-50 mg aliquots in sterile tubes at -80°C until required for downstream analysis.

2.1.2. Pituitary Tumours

Primary sporadic pituitary tumours were also investigated within this study. They comprised each of the major subtypes and were graded according to a modified Hardy classification (Bates *et al.*, 1997). The subtypes were as follows: growth hormone (GH)-secreting adenomas, corticotroph adenomas (CT), prolactinomas (PRL), and non-functioning adenomas (NF). Tumours were collected from patients during hypophysectomy. Adenoma subtype classification was on the basis of staining for mature hormone (GH, ACTH, FSH, LH and PRL but not for the α -subunit). The non-functional adenomas did not stain for mature hormones. All GH secreting adenomas were classified

as pure somatotrophinomas because they did not stain for any mature hormones other than GH.

Only those adenomas, in which tumour cells comprised at least 80% of the specimen, as determined at surgery and confirmed by neuropathological assessment, were used in the study. All tumours were freeze fractured using a biopulveriser as described for the normal pituitaries. Tumour tissues were obtained with informed consent and all studies performed with Regional Ethics Committee (South Birmingham Committee: REC reference number: 10/H1207/406) and institutional approval.

The primary sporadic human pituitary tumours used in this study comprised a discovery and an independent cohort (See Appendix 1). The discovery cohort comprised three each of the major adenoma subtypes. The investigation cohort comprised seven GH-secreting tumours all of which were grade 2 macroadenomas; seven corticotrophinomas (CT), four of which were grade 1 microadenomas and three were grade 2 macroadenomas; seven prolactinomas (PRL) five of which were grade 1 microadenomas and two were grade 2 macroadenomas and 13 non-functioning (NF) adenomas seven of which were grade 2 and six were grade 3 macroadenomas.

2.2. Cell culture methods

2.2.1. Pituitary Cell Lines:

1. AtT20 Cells

The AtT20/D16v-F2 (AtT20) murine pituitary adenoma cell line in the corticotroph lineage was purchased from the European collection of Cell Cultures (ECACC, Porton Down, Salisbury, UK). Cells were passage 13 upon purchase but were designated passage

0 for the purpose of these studies. Cells were not passaged more than 10 times for the studies described in this thesis.

II. GH3 Cells

The rodent pituitary adenoma cell line (GH3, ATCC code: CCL-82.1.) in the somato-lactotroph cell lineage were purchased from the American Type Culture Collection (Manassas, Virginia) Cells were passage 1 upon arrival and were designated passage 0 for the purpose of these studies. Cells were not passaged more than 10 times for the studies described in this thesis.

2.2.2. Growth Conditions

The GH3 and AtT20 cell lines were grown in Dulbecco's modified Eagle's Medium (DMEM) with 4500 mg/L of glucose, 0.584 g/L of L-Glutamine and 0.11 g/L of sodium pyruvate (Biosera, Ringmer, UK), and supplemented with 10% heat-inactivated fetal bovine serum (FBS, Biosera). Cells were incubated in a Thermo Scientific HEPA filtered single chamber water jacketed incubator, at 37°C in a humidified 5% CO₂ atmosphere.

2.2.3. Antibiotic Supplementation

To minimise the risk of bacterial infection the cell culture medium was supplemented with 4 µg/mL of aminoglycoside antibiotic Gentamycin (Sigma-Aldrich, Dorset, UK) and 2 µg/mL of antibiotic ampicillin (Sigma).

2.2.4. Detachment of adherent cells

Prior to experimental manipulation or routine passage of cell lines it was first necessary to detach and dissociate adherent cells from each other and from the base of culture flask. This was achieved using a solution of 1.7 M EDTA. Cell dissociation was facilitated by mechanical pipetting to further aid the release of the cells into solution before use. Cells were incubated for 5 minutes at room temperature and then following vigorous pipetting. Cells were neutralised by addition of 1 volume of DMEM containing 10% FBS. Resuspended cells were transferred to a sterile 15 mL polypropylene tube (Sarstedt, Leicestershire, UK) and centrifuged at 150 x g for 5 minutes at room temperature in MSE Mistral 2000 centrifuge (UK). The cell pellet was re-suspended in fresh medium by careful pipette mixing.

2.2.5. Culture Vessels

The type of vessels in which the cells were cultured depended on the experimental procedure being undertaken. For general sub-culturing purposes T25 filtered tissue culture flasks (Sarstedt, Leicester, UK) were used, whereas T75 filtered tissue culture flasks were used for the purpose of growing cells to sufficient density for cryogenic storage. A variety of tissue culture plates (96-well, 24-well, 12-well, 6-well; Sarstedt) were used for transfection and drug treatment studies.

2.2.6. Sub-culture

Sub-confluent (70-80%) cells were routinely sub-cultured in sterile conditions every three to four days. Cultured medium was removed and cells were washed with PBS, and then PBS-EDTA was added to cells and further diluted with media post detachment as

described above. Diluted cells were then transferred to a 15mL tube and centrifuged. Cells were then resuspended prior to counting with a haemocytometer. Cells were then diluted to 1×10^5 per mL and transferred to the appropriate culture vessel and incubated at 37°C. For routine culture cells were not grown beyond 10 passages and at this point, earlier passage and cryopreserved cells were revived from liquid nitrogen storage.

2.2.7. Cryopreservation of cells

For long term storage cells were cooled to sub-zero temperatures in a solution containing the cryopreservant dimethyl sulphoxide (DMSO). Cells were first grown to approximately 75% confluence in a T75 tissue culture flask and were then released from the adherent surface as described above and transferred to a 15mL polypropylene tube, and centrifuged at 150 x g for 5 minutes. Cell pellets were re-suspended in 900 μ L of FBS (90%) and 10% DMSO (Sigma), and transferred to a 2 mL cryovial (BD labs). Vials were placed into a “Mr Frosty” freezing container (Nalgene, Neerijse, Belgium) filled to the appropriate level with isopropanol (Sigma) and frozen overnight at -80°C. Frozen vials were then transferred to liquid nitrogen.

2.2.8. Reviving cryopreserved cells

Cells were revived from liquid nitrogen storage by rapid thawing to 37°C. Thawed cells were transferred to a 50 mL polypropylene tube and 20 mL of DMEM without FBS supplemented was added in a drop wise motion with constant swirling. Cells were then centrifuged 150 x g for 7 minutes and re-suspended in 6 mL of DMEM containing 10% FBS. Cells were then transferred to a T25 tissue culture flask and incubated overnight at 37°C. The following day medium was replaced in order to remove dead cells, whilst healthy cells were grown to sufficient density for sub-culturing purposes.

2.2.9. Cell Counting

Disposable haemocytometer chamber (Glasstic® Slide 10 with grids, Hyco Biomedical Inc, California, USA) were used to count cells. Cells were first stained with trypan blue to distinguish dead from live cells and suitable dilutions, such that between 20-50 cells, in each of the squares were enumerated to determine live-cell counts. The average number of cells per millilitre (cell/mL) was determined using the equation below.

$$\text{Total number of cells per mL} = \text{average count per mm}^2 \times 10^4 \times \text{Dilution factor}.$$

2.2.10. Incubation of cell lines with pharmacological agents

The epi-drugs in this study were 1-(β -D-ribofuranosyl)-1,2-dihydropyrimidin-2-one also known as Zebularine and Trichostatin A (Sigma, UK). Zebularine is a cytidine deaminase inhibitor and a DNA demethylating agent. TSA (Trichostatin A) is a histone deacetylase inhibitor (Cheng *et al.*, 2003; Cameron *et al.*, 1999).

Zebularine was prepared in DMSO and TSA was prepared in ethanol. Both drugs were stored at -20°C. Typically, drugs were stored as stock solutions: Zebularine (1 μ M) and TSA (10 ng/mL) and diluted prior to use.

24 hours prior to drug treatment, 2.5 mL of cells, at a concentration of 1×10^5 /mL were seeded into 6 well plates. Cells were then treated with the appropriate concentration of drug(s) (see figures and figure legends) every 24 hours for 48 hours. Post Zebularine and/or TSA challenges cells were harvested prior to extracting protein, nucleic acids and chromatin.

In some experiments, bromocriptine (BC) [2-Bromo- α -ergocryptine methanesulfonate salt] was used. This was obtained from Sigma. Stock solutions were

prepared in DMSO and PBS respectively and stored according to manufacturer's instructions.

2.3. Molecular biology methods

2.3.1. Isolation of genomic DNA from cell line and primary pituitary tissue

The extraction and purification of genomic DNA from cell lines and primary tissue (normal pituitaries and primary pituitary tumours) was performed using a standard lysis and phenol/chloroform procedure.

Lysis of tissue specimens: 3 mL of lysis buffer was added to 10-30 mg of tissue and incubated overnight at 56°C.

Lysis of cell line samples: Monolayers of cells ($1-3 \times 10^5$ cells) were washed once with sterile PBS and harvested into a PBS-EDTA solution. Cells were collected by centrifugation at 150 x g for 5 minutes and the pellets were resuspended in lysis buffer (see Appendix 2). Cell samples were incubated overnight at 37°C.

Extraction and precipitation of DNA: After overnight incubations samples were vortexed vigorously and an equal volume of phenol was added to lysates and mixed on a rotary shaker for 10 minutes. The samples were then centrifuged at 914 x g for 10 minutes. The upper aqueous phase was then transferred to a fresh tube, and an equal volume of chloroform added and mixed on a rotary shaker for 10 minutes, and then centrifuged 914 x g for 10 minutes at 4°C. The upper aqueous layer was transferred into a fresh tube and sodium acetate was added to achieve a final concentration of 0.3M. 100% ethanol was added and allowed to precipitate overnight at -20°C. This was then centrifuged at 150 x g for 25 minutes at 4°C. The supernatant was discarded and the pellet washed twice with

70% ethanol. After removal of the supernatant the pellet was left at room temperature to dry and resuspended in 50-150 μ L of double distilled RNase and DNase free water prior to quantification and storage at -20°C.

2.3.2. RNA extraction from cell lines and primary pituitary tissue

Total RNA was extracted using the guanidinium iso-thiocyanate-phenol-chloroform method previously described (Chomczynski and Sacchi, 1987). This is a chaotropic compound that denatures proteins including ribonucleases, whilst maintaining the integrity of RNA.

Isolation from cell lines: $1-3 \times 10^6$ cells were pelleted by centrifugation at 150 x g in a 15 mL polypropylene tube then washed in PBS, and re-suspended in 1mL of Guanidinium thiocyanate solution (see Appendix 2 for composition, concentration and storage of stock solution).

Isolation from tissue: 1mL of Guanidinium isothiocyanate lysis buffer was added to 10-30 mg of tissue and homogenised in a dounce homogeniser to a slurry like consistency. Lysates were then transferred to a 15 mL poly propylene tube.

200 μ L of 2M sodium acetate (pH 4.5) was added and mixed by vortexing. RNA was extracted from the solution by addition of equal volume of phenol and 800 μ L of 24:1 chloroform isoamyl-alcohol solution. Vigorous shaking ensured that the organic and aqueous phases were adequately mixed, and samples were incubated on ice for 30 minutes. The samples were centrifuged at 914 x g for 20 minutes at 4°C to separate the two phases; the upper aqueous phase was transferred to a fresh sterile tube. A double extraction was carried out by adding a further 800 μ L of chloroform isoamyl alcohol and shaking vigorously. Samples were incubated on ice for a further 30 minutes. Centrifugation was

carried out using the above settings and the upper aqueous phase was transferred to a fresh tube. RNA was precipitated by addition of equal volumes of 99% isopropanol and incubated overnight at -20°C. Precipitated RNA was pelleted by centrifugation at 914 x g for 30 minutes at 4°C. The supernatant was removed and the pellet was washed with 70% ethanol. Following re-centrifugation and removal of the supernatant the pellets were air dried to remove residual ethanol, and dissolved in 10-50 µL of sterile RNase and DNase free water (Sigma) depending on the size of the individual pellets. RNA samples were aliquoted out and stored at -80°C.

2.3.3. Purification of nucleic acid

During some experiments it was important to further purify the nucleic acid product. This was achieved using The GenElute[™] PCR Clean-up Kit which is designed to rapidly purify single and double stranded PCR amplification products (100-10 kb) whilst removing excess primers, nucleotides, DNA polymerase, oils and salts.

DNA is initially bound to a silica impregnated membrane within the spin column by the addition of a binding buffer. The bound DNA is washed with a wash buffer and the DNA is released from the membrane in elution buffer by centrifugation. Purified DNA is the suitable for several downstream applications including enzymatic digestions, conventional or automated sequencing, ligation, cloning and microarray analysis in this case prior to and following sodium bisulphite conversion.

2.3.4. Quantification of nucleic acid and assessment of purity

The quality of the nucleic acid isolated in any given RNA or DNA sample was estimated by spectrophotometric analysis using the NanoDrop 2000 (ND2000) (Thermo

Scientific, Nelson, UK). The NanoDrop is a spectrophotometer that measures micro volumes with a patent sample retention technology. It allows samples as small as 0.5µL to be measured without the need of cuvettes or capillaries.

The absorbance measurement made on the spectrophotometer will be of all molecules in the sample that absorb at the wavelength of interest. Since RNA, ssDNA and dsDNA all absorb at 260nm, they will contribute to the total absorbance of the sample so some samples will therefore require purification prior to measurement.

The ratio of the absorbance at 260nm and 280nm (260:280) is used to assess the purity of the DNA and RNA. A ratio of ~1.8 is typical for DNA while a ratio of ~ 2.0 is typical for high quality/purity of RNA. Lower ratios suggest the presence of proteins, phenol or other contaminants that absorb strongly at a wavelength of approximately 280nm.

2.3.5. Agarose gel electrophoresis

Electrophoresis refers to the separation of charged molecules in an electrical field. Molecules in a mixture are separated from each other on the basis of size, shape or charge. In this study a variety of electrophoresis methods were used to separate and visualise DNA and RNA molecules. To enable visualisation of the products from sqPCR agarose gel electrophoresis was employed. Low percentage (1-2%) agarose gels were used to separate DNA and RNA fragments. A typical 1% Agarose gel solution was prepared by melting 1 g of Agarose (Bioline, London UK) in 100 mL of 1X TAE buffer (diluted from 50X stock; see Appendix 2) using a microwave oven at high power for 2 minutes. The solution was then allowed to cool before addition of ethidium bromide (Sigma) to achieve a concentration of 0.5 µg/mL. The gel was then poured into a pre-assembled Bio-Rad gel

casting tray that had been sealed at each end using autoclave tape. A comb with the appropriately sized wells was immersed into the gel, which was then left to cool at room temperature for approximately 30 minutes in order for the gel to polymerise. The comb and autoclave tape were then removed, and the gel was immersed into a Bio-Rad resolving tank containing sufficient 1 X TAE buffer to cover the gel by 2-3 mm. 1 μ L of 6 X loading dye (see Appendix 2) was added to 5 μ L of nucleic acid samples, which were transferred to individual wells of the gel. Samples were electrophoresed from the negatively charged cathode to the positively charged anode for 45 minutes at 100V, alongside a GeneRuler 100 base pair DNA ladder (Fermentas, Yorkshire, UK), and nucleic acid fragments were detected and photographed by UV Transillumination using a Syngene gel documentation system (Cambridge, UK).

2.4. Whole genome amplification of bisulphite converted DNA

2.4.1. Sodium bisulphite modification of genomic DNA

Sodium bisulphite conversion of DNA was performed using a commercial kit. The kit used was EZ DNA Methylation-Gold™ Kit (ZYMO Research, Cambridge, UK). In this protocol 500 ng of genomic DNA is first treated with sodium bisulphite, essentially as first described in (Clark *et al.*, 1994). For efficient conversion genomic DNA was first denatured by incubating with conversion reagent and heated to 98°C for 10 minutes and then 64°C for a further 3 hours.

The converted DNA and M-binding buffer was added to a column (provided in the kit) that had been inserted into a 2 mL eppendorf tube (provided in the kit) and inverted several times. This was then centrifuged at full speed (914 x g) for 30 seconds. The flow-through was discarded and M wash-buffer was then added to the column. This was

centrifuged as described above. The flow-through was again discarded. The columns were centrifuged again at full speed to remove any residual wash buffer. The M-desulphonation buffer was then added to the column to desulphonate the DNA and left to incubate at room temperature for 20 minutes. The column was then centrifuged at full speed and all flow-through discarded. The desulphonated column-bound DNA was then washed twice by addition of M-wash buffer and centrifuged at full speed removing all flow-through. Finally the DNA was eluted into 21 μ L of elution buffer and then quantified on the Nanodrop ND2000 under the option of single stranded DNA. Samples were stored in -80°C after conversion.

2.4.2. Whole genome amplification

Whole genome amplification (WGA) was performed following a slightly modified version of the Zhang's primer extension preamplification (PEP) protocol (Zhang et al., 1992). 1 μ L bisulphite-treated DNA, corresponding to 20 ng of the BSC DNA, was used as template for whole genome amplification reaction. The reaction volume was increased to 25 μ L. To each of the WGA reactions 200 pmol of a 15-mer random primers (5'-NNNNNNNNNNNNNNNN-3' [where N is any of the four bases]) and 0.4 mM dNTPs were added. Amplification was performed for a total of fifty cycles. Each cycle including a denaturing step at 95°C for 1 minute, an annealing step at 37°C for 2 minutes, and an incubation step at 55°C for 4 minute. Ramping rate was set at 10°C/s . The WGA products were stored at 20°C .

2.5. DNA methylation analysis

2.5.1. Genome-wide DNA methylation analysis by Illumina BeadArray assay

I. Sample preparation and processing

For the discovery cohort, DNA was extracted from three of each of the major pituitary adenoma subtypes (NF, GH, PRL and CT) and from three NP. Total DNA (500 ng) from each of the samples was sodium bisulphite converted as described (see above). After conversion, DNA was eluted in 12 μ L elution buffer and 4 μ L converted DNA was used as template on the Infinium Methylation 27K Arrays (Illumina, San Diego, CA, USA) and was processed according to the manufacturer's recommendations. In these cases, Array hybridizations were performed under a collaborative agreement with Dr Charles Mein (Queen Mary's and London University Hospital (formerly Bart's), London, UK). This array examines the methylation status of 27578 CpG sites across 14496 genes. Data were collected using the Illumina BeadArray (Illumina) reader and analysed with GenomeStudio V2009.1 methylation module 1.1.1 (Illumina). This assigns a ' β value' that is a quantitative measure of methylation for each CpG site and ranges from 0 (no methylation) to 1.0 (100% methylation of both alleles).

II. Data analysis

The analysis of the initial, array generated data-set, was performed as part of collaboration with Dr Richard Emes (Reader and Associate Professor in Bioinformatics, University of Nottingham, UK). Subsequent to this analysis I was able to generate a data-set on the basis of the criteria described below. Initially, of the 27,578 sites on the array, we removed all probes targeted on X and Y chromosomes ($n = 1092$) and this reduced the

probe number to 26,486. Following discussions with my supervisor and with our collaborator (Dr Richard Emes) and to to reduce the number of non-variable sites from subsequent analyses, probes where β values in all samples were ≥ 0.8 or ≤ 0.2 were excluded ($n = 13,945$) from the analysis and this reduced the probe count to 12,541. From this dataset, I also eliminated all sites where one or more of the samples demonstrated i) detection P values of >0.05 (internal quality control) or ii) null (missing) β values. I took, after discussion, the decision to remove probes that failed in any one of the samples rather than excluding probe values for individual tumour samples as this was deemed a more robust way to identify probes that varied between tumours in a limited number of samples. This resulted in a final dataset of 10,667 probes for further analysis.

For the further analysis, that defined criteria for inclusion or exclusion of specific genes it was necessary for me to derive this data-set empirically. In these cases, these criteria are described in the results section of this thesis.

2.5.2. Bisulphite sequencing

1. Primer design for sodium bisulphite converted DNA

Primers, that were specific for sodium bisulphite treated DNA were designed for the amplification of bisulphite converted DNA. DNA sequences of genes of interest were obtained from Ensembl database (www.ensembl.org). The sequences were *in silico* bisulphite converted using the PyroMark Assay Design 2.0 software where the non-methylated cytosine would have been converted to uracil. The primer is only able to recognise this conversion if a thymidine (T) is incorporated into the design. Since methylated Cs are usually found in the context of a CpG they are not converted in a sequencing reaction and are read as C. Typically, for specificity and efficient

amplification primers were in the size range, 24-32 bases. Optimal amplicon sizes were between 150-300 bp. See Appendix 3 for Primer sequence for sodium bisulphite converted DNA.

II. PCR amplification

The polymerase chain reaction (PCR) is an *in vitro* method of DNA synthesis that allows particular regions of DNA (or cDNA in the case of RT-PCR) to be copied and amplified. The DNA template is first denatured by incubation at high temperature: then the temperature is lowered and two oligonucleotide primers that flank the DNA fragment to be amplified are annealed to their complementary sequences on opposite ends of the target sequence. The primers are then extended by Taq polymerase and the sequence between the primers is synthesised. Multiple rounds of denaturation, annealing and extension allow for specific amplification of the sequence of interest.

A touchdown PCR was performed using sodium bisulphite converted DNA or whole genome amplified DNA as template. It was necessary in some cases to perform nested PCR whereby the initial primary PCR product was subjected to a second round of PCR amplification.

III. Purification of DNA from agarose gels

Prior to T:A cloning of PCR amplicons it was necessary to first resolve them by agarose gel electrophoresis. This step ensured that the fragment corresponded to the right size and that it was separated from primer sequences. Fragments were, therefore, electrophoresed through a 2% agarose gel (as described in a previous section) and isolated and purified using a QIAquick Gel Extraction Kit (QIAGEN, West Sussex). The region of

agarose containing the DNA fragment was first visualised with a UV transilluminator and excised from the gel using a sterile scalpel blade. The gel fragment was then transferred to a 1.5 mL microcentrifuge tube and the volume of agarose estimated by measuring the mass of the excised fragment. 3 volumes (relative to the volume of the gel fragment) of Buffer QG was added and the agarose was melted by incubation at 50°C on a heat block for 10 minutes. A further 1 gel volume of isopropanol was added and the sample vortexed prior to addition of the DNA containing sample to a column provided within the kit. The sample was the subject centrifugation and the flow through discarded. This step was repeated. After further wash steps using Buffer PE the column was put into in a 2 mL collection tube and elution buffer EB used to release bound DNA as described in the manufacturer's protocol.

IV. Ligation

The ability of Taq DNA polymerase to extend PCR amplicons by the template independent addition of a single adenine (A) to the 3'-end, enables the rapid and efficient ligation of PCR products into a T:A cloning vector, the pGEM T-Easy vector (Promega, Southampton, UK, see Appendix 4). The vector is a pre-linearized with a single 3' – terminal thymidine at either ends. The overhang at the insertion site greatly increases efficiency of ligation of PCR products. PCR reactions were first performed in duplicate and then combined in a single tube. 6uL of this reaction mix was resolved on a 1% gel to confirm amplicon size. The remaining reaction was then cleaned on a Gene Elute PCR cleaning column (Sigma-see above). The quantity of the purified PCR product was determined on the Nano-drop (OD 260nm). An appropriate quantity (see calculation below) was then used for ligation into the pGEM vector at a 6:1 insert:vector molar ratio. This was calculated according to the following equation.

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert:vector molar ratio} = \text{ng of insert}$$

50ng of the 3000bp size pGEM vector was typically used in a ligation reaction; therefore at a 6:1 insert:vector molar ratio, approximately 5 ng per 100 bases of insert was used. To this reaction 400 U of T4 DNA ligase (Promega) and 4.5 uL of 10X T4 DNA ligase buffer was added. Ligation reactions were incubated at 4°C for 16 hours.

Ligation reaction:

Ligation Reagents	1 reaction (μL)
10 X T4 Ligase buffer	4.5
T4 Easy Vector	0.7
PCR product	1-4
T4 DNA Ligase	4
H2O (10 μL Total)	(~10)

Table 2.1. Ligation Reaction. Ligation reagents for one reaction.

V. Transformation

8 μL of the total 10 μL reaction was added to a 50 μL volume of competent cells (See appendix III on growing chemically competent cells) in a sterile 0.5 mL microcentrifuge tube, and incubated on ice for 20minutes. The microcentrifuge tube was gently flicked to mix the DNA and cells. The cells were then heat shocked for 2 minutes at 42°C in a thermal cycler, followed by a further incubation on ice for 2 minutes. The transformation mixture was then transferred to a sterile 15 mL polypropylene tube followed by the addition of 1mL sterile LB medium. Samples were incubated at 37°C for 2.5 hours in an orbital shaker with vigorous shaking.

VI. Spread plating

350µL of the bacterial culture was spread onto a LB-agar plate containing 60µg/mL of ampicillin using a sterilised spreader. In addition, 0.1M IPTG (isopropyl β-D-1-thiogalactopyranoside) and 50 µL/mL Xgal (5-bromo-4-chloro-3-indolyl β-D-galactopyranonide) were added to agar plates if TA cloning was being carried out, as this allowed for blue-white colony selection. Plates were incubated inverted at 37°C for 16 hours.

VII. Screening

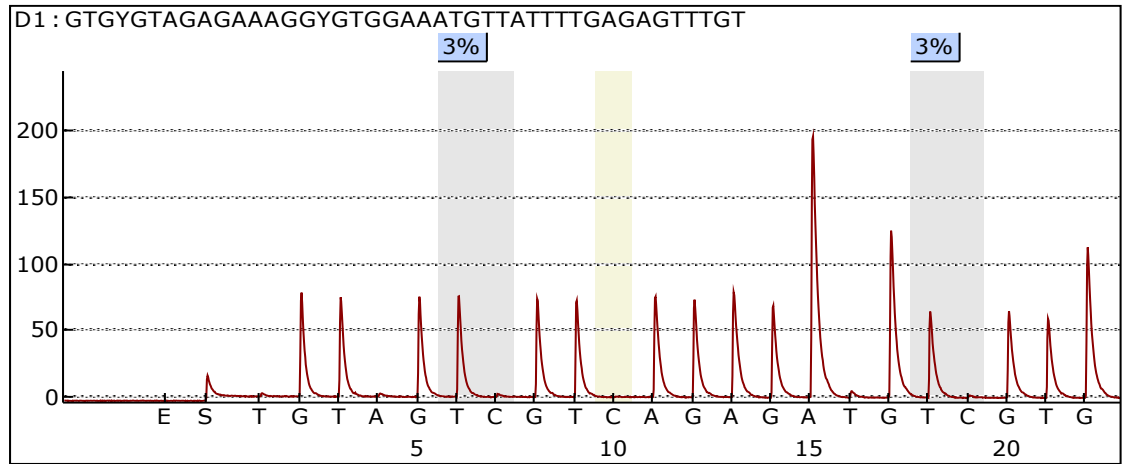
All vectors used during this investigation contain the ampicillin resistance gene, AMP^r, which meant that inclusion of ampicillin in the LB agar enabled positive selection of transformed colonies. In addition, successfully ligated PCR inserts into the T-Easy vector were screened by a simple colour selection procedure due to the presence of IPTG and Xgal in the LB agar. This substrate is catalysed by the enzyme β-galactosidase (product of the lacZ gene), yielding colonies that appear a distinctive blue colour. As the incorporation of a PCR product into the pGEM vector destroys the lacZ gene, β-galactosidase is not produced by successfully transformed cells, which therefore produces white colonies.

To confirm that successfully transformed cells contained an insert, it was necessary to carry out sequencing analysis. Colonies were picked off the agar plate using a sterile pipette tip and lysed in 50 µL of ddH₂O. This was done by heating at 96°C and then centrifuging at 914 x g for 10 minutes. 5 µL of the lysate was subjected to sqPCR amplification using primers that flank the cloning site SP6 reverse and T7 forward. PCR products were purified using the Gene Elute PCR Purification kit (Sigma) and sent to be

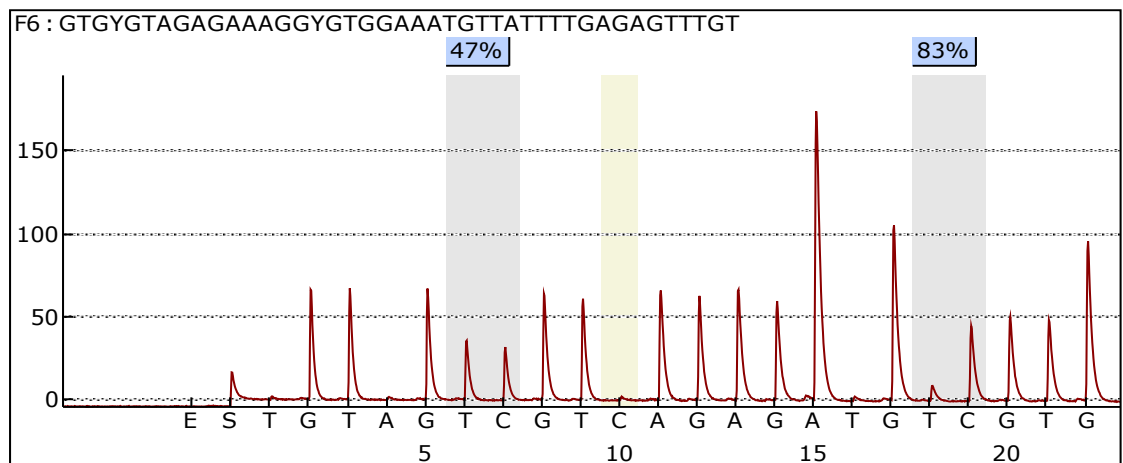
commercially sequenced (gene service, Cambridge UK) using the T7 forward primer. Typically, 4-5 molecules that were isolated from individual bacterial colonies were sequenced from each of the sample.

2.5.3. Pyrosequencing

Pyrosequencing, of sodium bisulphite converted DNA, is a sequence based analysis and can be used to accurately quantify the level of methylation at CpG sites. In this way it provides data on the individual level of methylation in the context of the underlying DNA sequence (Colella *et al.*, 2003). Pyrosequencing uses a single-stranded DNA templates for synthesis of a complementary strands and this is achieved through the sequential addition of the four nucleotides (A, T, G, and C), where every successful nucleotide incorporation releasing pyrophosphate (PPi). PPi is converted to ATP by the enzyme sulfurylase and subsequently emits a light signal catalysed by the enzyme luciferase which is proportional to the number of incorporations. A pyrogram is generated with peak heights representing the intensity of light emitted from successful nucleotide incorporations informing on the sequence content in the analysed fraction. The degree of methylation is calculated from the pick heights of C and T as demonstrated in figure 2.1.



(A)



(B)

Figure 2.1. Representative pyrosequencing pyrogram.

A sample read out of the program generated by the Pyro Q-CpG program of the EFEMP1 gene in a human post mortem pituitary (A) and a non-functioning pituitary tumour (B). The top left of each pyrogram shows the sequence analysed, where each of the individual CpG dyads are shown as YG. The sequence shown on the x axis is the nucleotide dispensation order, where the ratio between the T and C peak heights (wide grey vertical highlight) defines the percentage methylation (shown on the y axis). As internal control the efficiency of bisulphite conversion was determine from a cytosine (C) that was not in the context of a CpG dyad. The internal C is highlighted as a narrow grey vertical highlight.

I. Assay design

For each gene, the methylation status of its promoter region-associated CpG island in normal pituitaries, cell lines, and primary human pituitary tumours was determined using primers that encompassed CpG dinucleotides within *bona fide* CpG islands. In these cases the islands were upstream of the transcription start site and frequently extended into the first exon of the gene. The primers used to examine the CpG island in each of the species, the number of CpG interrogated by pyrosequencing, and the sequence to analyse of each assay are shown in Appendix 3. These primers were designed to recognize sodium bisulphite converted DNA. For each gene, DNA sequences were downloaded from the Ensembl genome browser (<http://ensembl.org>). CpG islands were identified using online tool (<http://cpgislands.com>) and imported into PyroMark Assay Design 2.0 software for primer design of sodium bisulphite-converted DNA (Qiagen, Crawley, UK). In some cases the sequence analysed in human pituitary and adenomas encompassed and interrogated 1 or 2 CpG sites which had also been interrogated in BeadArray analysis.

II. PCR amplification

2 μ L sodium bisulphite-converted and whole genome amplified DNA were used as template in the first round of a nested PCR reactions. The product of the first round was diluted 100-fold in dH₂O and 2 μ L used as template in a second round nested PCR reaction. In the second round of each PCR reaction, 2 μ L of template DNA, 1 unit of *Taq* DNA polymerase, 100 μ M of dNTPs, and 10 pmol of each primer (see Appendix 3) in total volume of 25 μ L. The same cycling condition were used for each round of amplification. After initial denaturation at 98°C for 10 min, touch-down cycling was employed for 14 further cycles, where, in each successive cycle the temperature was

touched down by 0.5°C. The subsequent 35 amplification cycles include denaturing at 98°C for 30 s, annealing at 55°C for 30 s and elongation at 72°C for 30 s. Additional elongation was set at 72°C for 10 min. Size and quantity of PCR products derived from the second round of amplification were determined on a semi-quantitative basis by visualization on agarose gel electrophoresis.

III. Pyrosequencing

20 µL of PCR products were mixed with 50 µL of streptavidin beads in binding buffer (Biotage, Upsala, Sweden) (10 mM Tris, 2 M NaCl, 1mM EDTA, 0.1% Tween 20, pH 7.6) at room temperature on a rotary shaker for at least 10 minutes. The biotinylated strand formed a complex with the streptavidin beads due to the high affinity of biotin for avidin. During this binding reaction, a master mix of the annealing buffer, comprising each of the sequencing oligonucleotide (see Appendix 3) was prepared (0.5 µL of 100 pmol sequencing primer and 39.5 µL of annealing buffer) and 40 µL was pipetted into each well of a 96 well pyrosequencing plate. Once the plate was prepared the Biotage pyrosequencer (PSQ 96MA) (Biotage, Upsala, Sweden) was programed with the sequence for analysis and the dispensation order for the run. The PCR product/bead mixture was transferred from the shaker and the capture of the single-stranded and biotin-labelled DNA achieved using a Pyromark Q96 ‘Vacuum Prep’ Workstation (Biotage, Upsala, Sweden). The workstation aspirated the binding mixture which filtered the streptavidin-biotinylated single strand from the non-biotinylated single strand and removed all remnants of the PCR reaction by washing first with 70% ethanol, strand separation with 0.2 M NaOH and neutralisation with washing buffer (10 mM Tris pH7.6, adjusted with acetic acid). The washed biotinylated strands were then released into the prepared pyrosequencing plate. The plate was placed on a heating block at 80°C for 2 minutes to allow sequencing primers

to anneal to the target and then loaded into the pyrosequencer. The pyrosequencer cartridge contained the Pyro Gold sequence analysis reagent, provides each nucleotide, substrate, and the enzyme mix.

Pyrosequencing was performed using a PSQ96A Pyrosequencer using Pyro Q CpG Software (version 1.0.9; Qiagen). For each set of independent bisulphite modification reactions, a control DNA sample was run to ensure efficient sodium bisulphite-mediated conversion. In addition, another bisulphite converted control was generated by adding of an additional C/T incorporation at non-CpG site in the sequence subject to-analysis (see Figure 2.1). The dispensation order (see Appendix 3) instructs the pyrosequencer to dispense each nucleotide based on the sequence order dictated for the analysis and this can be modified to dispense more than one nucleotide at a particular site in order to determine which of the nucleotide is incorporated. In this instance the dispensation order was modified to dispense a C after a T nucleotide at each CpG site to provide information on the level of incorporation of either of these nucleotides. This gives an accurate measure of methylation which is directly related to successful C incorporations (see Figure 2.1). The results were analysed using the Q-CpG (Biotage) program and displayed as a pyrogram with each CpG site assigned percentage cytosine content and the mean of methylation level across all CpG sites interrogated in each assay.

2.6. Gene expression analysis

2.6.1. cDNA synthesis

First strand cDNA synthesis was carried out from 1 µg of good quality total RNA prepared as described above. For the synthesis, 250 ng of random primers (Promega, Southampton, UK) was added and incubated at 70°C for 5 minutes to melt and denature

any secondary structure. Samples were snap frozen before addition of 2.5 μ L of 5 μ M dNTP mixture and 200 units of M-MLV RT (Promega) and a reverse transcriptase buffer added to yield a final reaction volume of 25 μ L. Samples were then incubated at 37°C for 1 hour and stored at -80°C until required for RT-PCR.

2.6.2. Primer design

Primer 3 online software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) was used to design all PCR primers used throughout this project, other than those used for PCR amplification of sodium bisulphite converted DNA and described in the preceding section. Genomic and cDNA sequences were obtained from the Ensemble genome browser (www.ensembl.org). Primer sequences with minimum self-complementarity, especially 3' self-complementarity were chosen to avoid and mitigate against primer dimer amplification. In addition, primers yielding amplicons within a size range 100-300 bp were preferred, as these are known to produce more reliable results in quantitative analytical procedures.

See Appendix 3 for a list of all primer sequence and amplicon sizes used throughout the study. All primers used throughout the study were purchased from Biomers and Invitrogen.

2.6.3. qRT-PCR

Real time quantitative qPCR allows quantification of PCR products in “real time” during each successive PCR cycle in the course of the reaction. Such reactions are carried out in a thermocycler that permits measurement of the fluorescent detector molecule such as the intercalating dye SYBR green, which fluoresces upon incorporation into double

stranded DNA. Increased fluorescence occurs as more dye is incorporated into the DNA as each cycle of the PCR reaction. Detection of fluorescence at each successive cycle allows an amplification plot to be generated. As well as decreasing post-processing steps this methodology minimises experimental error since SYBR Green is not sequence dependant and (unlike other qPCR probes for example TaqMan), it can be used for any reaction.

Quantitative PCR amplification was performed using a Stratagene Mx3005P thermal cycler (Agilent, Cheshire, UK). Reactions were prepared containing 1X Brilliant III SYBR Green QPCR mastermix (Agilent), 0.5 μ L 20 pmol of forward and reverse primer and ddH₂O. 11.5uL of the reaction mixture was added to individual wells of 96 well plate or to individual wells of 8 strip tubes (Agilent) and 1 uL of either cDNA template or no template control was added to each well. All samples were analysed in triplicate to account for technical variation. The following reaction conditions were carried out, which comprised activation of hot start Taq (96°C for 3 minutes) followed by 40 cycles (96°C 30 seconds followed by an annealing and elongation step for 30 seconds). A dissociation curve was included at the end of each reaction to ensure that specific amplification has been achieved.

This method is relative, as it requires normalisation to an endogenous control gene. In the studies described in this report the endogenous controls were Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) or Porphobilinogen deaminase (PBGD) (see appendix 2 for primer sequences). There are two methods by which qRT-PCR can be performed: the relative standard curve and the $2^{-\Delta\Delta CT}$ method. The relative standard curve method uses serial dilutions of cDNA samples, of known concentrations, from which unknown samples are quantified. This is achieved by amplification of the standards alongside the unknowns. It is possible to derive a standard curve for both the target gene and the endogenous control, thereby permitting relative levels of expression to be determined. In these types of

experiments it is important to include a standard curve on every plate analysed to counteract potential variability between PCR runs. This method requires minimal validation because the PCR efficiencies of the target and endogenous control do not have to be equivalent as different efficiencies are accounted for by the standard curve.

For samples analysed by the 2- $\Delta\Delta$ CT method the comparative differences between the gene of interest and the endogenous control are calculated by a mathematical formula, and as such a standard curve is not required. First, the difference in the cycle threshold (CT) of the gene of interest and endogenous control is calculated (the Δ CT). Next, subtraction of the control Δ CT from the treated Δ CT yields the $\Delta\Delta$ CT. The negative value of subtraction, the $-\Delta\Delta$ CT, is used as the exponent of 2 in the equation and represents the difference in “corrected” number of cycles to threshold. For all genes analysed using this method it is necessary to ensure the template amplification for the target gene occurs at the same efficiency as for the endogenous control. Therefore, during optimisation experiments, the standard curve analyses are carried out as described above, however, the difference in CT between endogenous control and gene of interest are plotted against log (base 10) input across the dilution range. Regression slopes <0.1 are considered to be acceptable for 2- $\Delta\Delta$ CT analyses.

2.7. Histone modification analysis

Chromatin immunoprecipitation assay (ChIP) is a powerful technique allowing analysis of protein modifications within specific genomic regions. ChIP is used to determine changes to epigenetic signatures, chromatin remodelling and transcription regulators that are recruited to specific genomic sites. Although ChIP is a versatile and powerful tool it is technically demanding and for the studies described in this thesis the Active Motifs ChIP-IT Express Enzymatic Kit (Rixensart, Belgium) was employed. There

are five main stages to the procedure. In brief intact cells are fixed using formaldehyde, which crosslinks and preserves protein and DNA interactions. The DNA is then sheared using an enzymatic digestion cocktail containing micrococcal nuclease, into small uniform fragments. Specific protein and DNA complexes are immunoprecipitated using antibodies directed against the DNA-binding protein of interest (see below for more detail). Once antibodies have bound the enriched fractions are pulled down with the aid of magnetic beads coated in protein G which has a high binding affinity for the antibody. Following immunoprecipitation cross linking is reversed, the proteins are removed by treatment with proteinase K and DNA is recovered. The DNA is then analysed to determine which DNA fragments were bound to the protein modification of interest. In this study this was achieved through qRT-PCR. The steps of ChIP protocol are illustrated in figure 2.2.

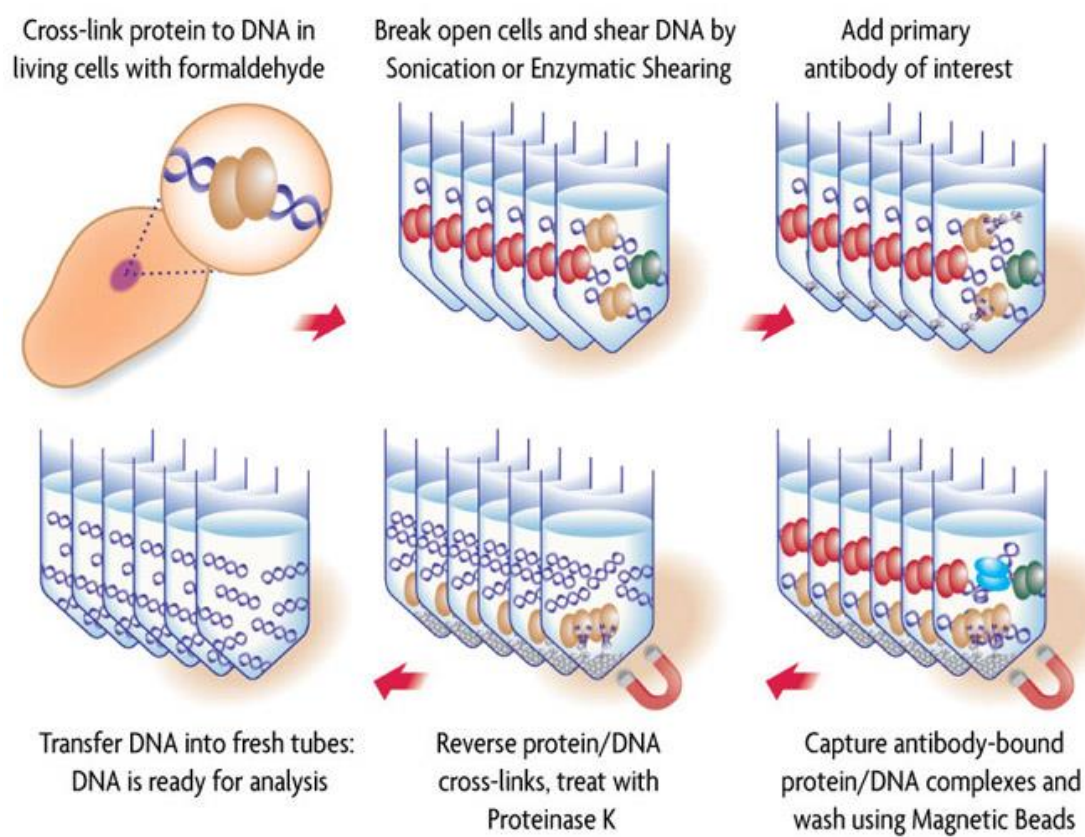


Figure 2.2. Schematic of chromatin Immunoprecipitation using ChIP-IT™ Express.

Adapted from ActiveMotif Epigenetic Products and Services 2013 Brochure

2.7.1. Cell collection and DNA-protein crosslinking

Cells were first grown to 70-80% confluency (approximately 1-1.5 million cells) in a T75 flask. Cells were then detached from the surface of the flask using PBS-EDTA and transferred to a siliconised 50 mL tube and centrifuged at 150 x g for 5 minutes. The supernatant was discarded and the pellet washed with ice cold PBS. This was again centrifuged. A solution of DMEM media containing a final concentration of 1% formaldehyde was added to cells and the cell pellets were disassociated by pipetting and incubated on a rotary shaker for 5 minutes. For tissue samples, the 1% formaldehyde-

DMEM solution was added to the tissue and homogenised in a dounce homogeniser. 500 μ L of the glycine solution, supplied with the kit, was added at room temperature and rotated on the rotary shaker for 5 minutes to stop excessive crosslinking and centrifuged as described above.

2.7.2. Cell lysis and enzymatic chromatin shearing

After the fixation step described above, the cells were first lysed in the lysis buffer supplemented with PIC (Protease Inhibitor Cocktail) and PMSF (Phenylmethanesulfonyl fluoride) for 30 minutes on ice. During this incubation the enzymatic shearing cocktail was prepared by diluting the supplied enzymatic shearing cocktail (2×10^4 U/ml) to 1:100 with a 50% glycerol solution. For the chromatin shearing step, and to generate fragments in the size range 150-1500 bp, it was necessary to perform enzymatic shearing in the presence of micrococcal nuclease. For these fragment sizes it was first necessary to optimise shearing in time course experiments. Therefore, the lysis solution together with the cells were transferred to an ice cold dounce homogeniser and then gently homogenised for 10 strokes on ice to aid nucleic acid release. After transfer to a siliconised microcentrifuge tube and centrifugation ($914 \times g$ for 10 minutes at 4°C) the supernatant was discarded and the pellet resuspended in warm digestion buffer supplemented with PIC and PMSF and incubated at 37°C for 5 minutes. 17 μ L of the enzymatic shearing cocktail was added to the pre-warmed nuclei and subjected to gentle vortexing and incubated for 12 minutes at 37°C with intermittent vortexing. The enzyme activity was stopped by addition of 7 μ L of ice cold EDTA and incubated on ice for 10 minutes. All samples were then centrifuged at $914 \times g$ for 10 minutes at 4°C and the sheared chromatin supernatant collected. This could be stored at -20°C for later use. In this study, optimal enzymatic chromatin shearing generated bands and smearing between 150-1500 bp as shown in figure 2.3

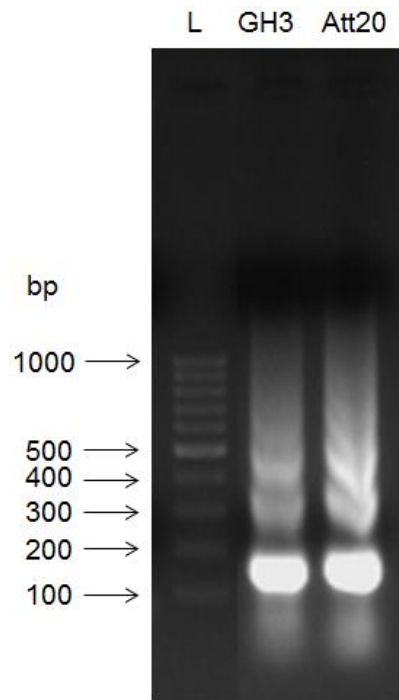


Figure 2.3. Gel analysis of optimal enzymatic shearing.

L: 100 bp ladder; GH3 and AtT20: sheared chromatin derived from GH3 and AtT20 cells, respectively

2.7.3. Immunoprecipitation

The antibodies used for the described studies were purchased from Abcam (Cambridge, UK). The antibody used to detect active genes recognises the H3K9Ac (histone 3 lysine 9 acetylation) modification (Jenuwein and Allis, 2001). The antibody used to detect silenced genes recognises the H3K27Me3 (histone 3, lysine 27 trimethylation) modification (Cao *et al.*, 2002a).

After thawing the samples generated in step 2 (see above), 10 μ L of the chromatin was set aside and represented input DNA. For the samples subject to immunoprecipitation, 60 μ L of sheared chromatin was added to siliconised tubes together with the following reagents; 25 μ L protein G beads, 10 μ L Chip Buffer, 1 μ L PIC water and modification specific antibody and transferred to a rotary shaker for 4 hours at 4°C. The protein G beads

were magnetic beads coated with protein G. The protein G has a binding affinity for the antibody heavy chains and was, therefore, able to pull down the antibody in a complex with the specific histone modification. The DNA-histone complexes were immunoprecipitated and washed. Immunoprecipitation was achieved by placing the tubes on a magnetic stand to pellet the beads and the supernatant was removed without disturbing the beads. The beads containing the complex were washed with ChIP buffers 1 and then 2 several times and pelleted using the magnetic stand.

2.7.4. Reversal of Crosslinking and DNA isolation.

Following the wash steps, the chromatin was eluted from the antibody that was in turn bound to the protein G beads. This was achieved by addition of 50 μ L elution buffer and incubating for 15 minutes with intermittent pipette mixing at room temperature. The reverse cross linking buffer was added (50 μ L) and tubes immediately transferred to the magnetic stand to allow the beads to pellet. The supernatant (100 μ L at this point), which contained the chromatin was transferred to a fresh tube.

The immunoprecipitated chromatin and the input chromatin (see step 1) were treated separately. 88 μ L of ChIP buffer was added to the 10 μ L of input sample. 2 μ L sodium chloride solution was added to both input and immunoprecipitated samples and heated to 95°C for 15 minutes in a thermal cycler. Proteinase K (2 μ L) was then added to both immunoprecipitated chromatin and Input chromatin and incubated at 37°C for 1 hour. This step effectively digests proteins including the histones that were in complex with the DNA. After the incubation period proteinase K stop solution (2 μ L) was added to the reaction. The DNA within this solution was then concentrated by binding onto PCR GeneElute columns (as described in a previous section).

2.7.5. Quantitative PCR and data analysis.

Primers were designed to amplify a region within 300 bp of the known or presumed transcription start site (TSS) (Roh *et al.*, 2006). In cases of human and mouse, TSSs of *EFEMP1/Efemp1* genes were obtained from The Eukaryotic Promoter Database (EPD) available online at <http://epd.vital-it.ch/>. This database was considered at the highest priority for selecting TSS because it provides experimentally validated promoters (Dreos *et al.*, 2013). However, EPD does not include rat genome. Therefore, the TSS of rat *Efemp1* was obtained from University of California, Santa Cruz (UCSC) Genome Browser available at <http://genome.ucsc.edu>. After determining TSS, the sequence of -150 to +150 relative to TSS was retrieved and served as template for primer design using Primer 3 online tool <http://bioinfo.ut.ee/primer3-0.4.0/>.

The input DNA was further diluted 1:10 using ddH₂O. A qRT-PCR was set up using the DNA isolated by immunoprecipitation of the sheared chromatin and also the Input DNA as a control sample for 50 cycles. The method used to normalise the ChIP-qPCR data is the percentage input method. It is important to normalise against the input as the input sample represents the total chromatin pre-enrichment. There are several dilution steps subject to immunoprecipitation it is therefore necessary to use a dilution correction factor for the input. This allows you to calculate equal concentrations of the immunoprecipitated sample and the input. It allows comparison of “like with like” samples. The input sample represented 1/60 of the sample used for immunoprecipitation and therefore necessary to adjust this to 60%. The correction factor of this is to subtract 5.9 Ct values (The starting material is 1/60, a dilution factor of 60 needed to be taken into consideration, this is 5.9 cycles and was found by log₂ of 60. This was then subtracted from the Ct values of the input as demonstrated in figure 2.4).

The following equation was used to calculate the percentage enrichment relative to the input $100 \times 2^{(\text{adjusted input} - \text{CT IP})}$ see figure 2.4 below for a breakdown of this calculation.

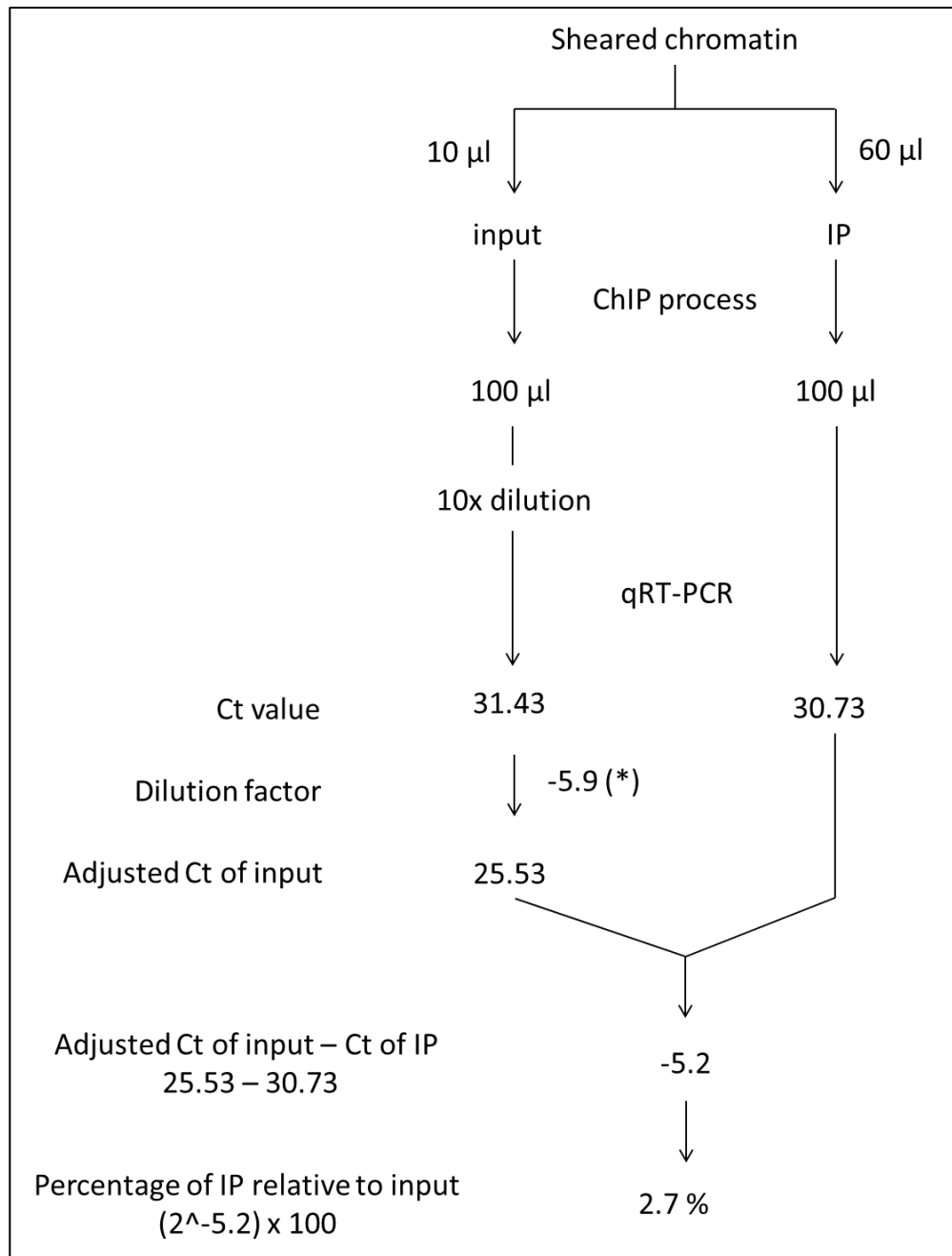


Figure 2.4. ChIP calculations.

A detailed example of the ChIP calculations used within this study. IP is the immunoprecipitated sample.

(*) 5.9 cycles was found by log2 of 60 fold dilution

2.8. Functional analysis

2.8.1. Subcloning of target gene into pTUNE vector

The complete coding sequence of human *EFEMP1* (GenBank accession no. NM_004105) was purchased as a cDNA clone (OriGene). The entry vector containing *EFEMP1* and the empty pTUNE vector (see Appendix 4) were transformed into *E. coli* strain DH5 α by heat shock as described previously in this thesis and spread onto LB-agar plate containing 60 μ g/mL kanamycin and 60 μ g/mL ampicillin, respectively. Individual colonies were grown at 37°C overnight in LB media supplemented with appropriate selective antibiotics. *E. coli* cells were pelleted by centrifugation at 4000 RPM and plasmids were extracted by QIAprep Spin Miniprep Kit (Qiagen) following manufacturer's protocol.

1 μ g of *EFEMP1* entry vector and empty pTUNE vector were digested by *KpnI* and *XhoI* restriction enzymes in 1X Tango buffer (Thermo Scientific) at 37°C for 90 minutes. Digested products were purified, ligated, and transformed as described in 2.5.2 section. Transformed *E. coli* cells were plated on the surface of LB plate supplemented with 60 μ g/mL ampicillin. Individual colonies were screened by PCR using human *EFEMP1* qRT-PCR primers (Appendix 3). Positive clones were confirmed by digestion with *KpnI* and *XhoI* restriction enzymes where bands correspond to *EFEMP1* and linearized pTUNE vector (1482 bp and 12.1 kb, respectively) were visualized. One single clone of pTUNE-*EFEMP1* (clone number 2) was selected for the subsequent experiments.

2.8.2. Stable transfection of rat pituitary cell line GH3

Lipofection, was the technique of choice and was used to introduce genetic material into cells via liposome mediated endocytosis. Liposomes are artificially prepared vesicles that are composed of a phospholipid bilayer, a property that allows them to easily fuse with the mammalian cell membrane. This technique is efficient and is capable of transfecting most types of nucleic acids into a wide range of cell types. This technique is a simple, reproducible and in general shows minimal toxicity. For the studies described in this thesis this technique was used exclusively for the introduction of inducible expression vector pTUNE into GH3 cells.

The day before transfection cells were seeded at a density of 2×10^5 cells per well in a 6 well plate such that the cells would be at approximately 50% confluence at the time of transfection. For each transfection two separate solutions were prepared in sterile 1.5 ml tubes. The first solution contained 15 μ L of Lipofectin and 85 μ L of serum free DMEM culture medium and was left at RT for 45 minutes. The second solution contained 10 μ L of DNA (at 1 μ g/ μ L) and 90 μ L of serum free DMEM culture medium. The two solutions were combined, mixed by inverting and incubated at RT for 15 minutes. During the course of this incubation the media was removed from the cells and cells washed twice with 2 mL of serum free media. Following incubation 800 μ L of serum free media was added to the combined solution and the 1 mL DNA/lipofectin/media mixture was dispensed onto the cells in drop wise manner. Cells were then incubated in cell culture incubator at 37°C for 6 hours. The transfection solution was then replaced by 2 mL of complete media and the cells allowed to proliferate for 48 hours. Cells were then detached and diluted 1 in 5 into separate chambers of a 6 well plate. Selective agent, geneticin, was added to the cells at 600 μ g/mL and cells grown for 14 days with washing and feeding every two days. Cells

harbouring the inducible vector, on the basis of their antibiotic resistance were harvested and used for subsequent analysis.

2.8.3. Characterization of stable transfected cells

Inducible expression of *EFEMP1* was determined by treating cells with a range of IPTG concentrations and measuring expression of *EFEMP1* at transcript level by qRT-PCR. Cells were seeded at density of 1×10^5 cells/mL into 25 cm² culture flasks in total volume of 5 mL of media and induced with 2.5 μ M, 5 μ M, 25 μ M, 100 μ M, and 500 μ M of IPTG for 24 hours. Cells were then harvested and used for total RNA extraction following by cDNA synthesis. qRT-PCR was performed using transcript specific primer for the *EFEMP1* and rat *Pbgd* genes (see Appendix 3).

2.8.4. Cell proliferation analysis by growth curve

The effect of induced *EFEMP1* expression was first assessed in stable transfected clones (GH3-EFEMP1) relative to cells harbouring an empty vector control (GH3-Cont). GH3 cells were seeded into six-well plates (Corning) at a density of 2.5×10^5 per well supplemented with 600 μ g/ml of Geneticin and induced with 100 μ M IPTG or vehicle control. Media were replenished at 2 day intervals. Triplicate individual flasks were sacrificed at each time point (2 day) and cells were counted and their viability determined by trypan blue exclusion. The experiment was repeated three times.

2.8.5. Apoptosis analysis by caspase 3/7 assay

Activated caspases were detected using Caspase-GloTM KIT (Promega) that specifically detects either of the initiator caspases (Caspase-8 or Caspase-9) or the

terminal caspases (Caspase 3 and 7). Caspase-Glo assay is a luminescent assay that utilizes the proluminescent caspase substrate and a thermostable luciferase in a single reagent optimized for caspase activity, luciferase activity, and cell lysis. Addition of Caspase-Glo reagent to the cells results in cell lysis, followed by caspase cleavage of the substrate to generate aminoluciferin. The liberated aminoluciferin is consumed by the luciferase enzyme generating a “glow-type” luminescent signal. The signal is proportional to the specific caspase activity. This assay system enables detection of caspase activity in a multiwell plate format using cells (treated or untreated) in culture.

Cells were seeded into 96 well plates at a density of 1×10^5 cells/ml and incubated overnight in growth media. For various duration of time, semi-confluent cultures were challenged with bromocriptine, shown previously to induce apoptosis in the cells (Rowther *et al.*, 2010), but in this case in the absence of presence of inducer IPTG. Duplicate sets of the cells treatments were carried out such that one set was used for caspase activity while the second one was used for the cell number correction (see 2.8.6).

The assay reagents which include caspase-glo substrate, caspase-glo buffer, and MG-132 (proteasome inhibitor, optimal for caspase-8 and 9 assay) were combined to make a ready to use assay reagent. 50 μ L of the culture media was aspirated from each well to be analysed for caspase activity (originally each well contained 100 μ L culture media). The ready-to-use reagent was then added to each well of the 96 well plate containing the treated or untreated cells in 1:1 ratio (50 μ L of culture media : 50 μ L of ready-to-use caspase-glo reagent). The plates were rocked sideways to mix the contents of the wells and incubated (covered by an aluminium foil) at room temperature for 45 minutes. Assay was performed for each type of sample in triplicate wells. At the end of the incubation, 75 μ L of the content of the wells were transferred to a detection tube (Sarstedt) and the luminescence was recorded using a tube luminometer (Lumat LB9507, Berthold Technologies,

Hertfordshire, UK) set to read for 10 seconds. Data was analysed after correcting for the cell numbers measured by crystal violet staining.

2.8.6. Cell correction using crystal violet staining

In order to normalize the readings obtained from caspase assay to the number of cells, cell correction was carried out. Crystal violet (CV) was used to stain the cells. CV is an intense stain that binds to the cell nuclei and gives an OD reading at 595 nm which is proportional to cell number.

Semi-confluent cultures were challenged with bromocriptine in the absence or presence of IPTG. Following challenge the cells were fixed by replacing the growth media with 100 μ L of 4% formaldehyde (Sigma) in PBS (Sigma). To minimize the escape of formaldehyde vapours, a piece of parafilm was placed over the plate and covered with the lid. The covered plate was incubated at room temperature for 20 minutes. At the end of the incubation, formaldehyde solution was removed and the cells were washed 3 times with 200 μ L of 1X PBS buffer. Each wash step was performed for 5 minutes with gentle shaking. Washed plate was air-dried for 10 minutes and used for CV staining. 100 μ L of 1% SDS solution (Sigma) was added to each well and incubated at room temperature for 30 minutes. The wells were washed again and air-dried before adding 100 μ L of 1% SDS solution (Sigma) to each well. The plate was incubated on a shaker for an hour at room temperature. The absorbance was determined on a spectrophotometer (Dynatech MR5000) at 595 nm. Occasionally when the signals obtained for certain samples were greater than the range of the spectrophotometer, the signal was reduced by removing some (e.g. 50 μ L) of the liquid from each well and replacing with an equivalent volume of ddH₂O.

Section C: Results and Discussions

Chapter 3: Whole genome amplification of bisulphite converted DNA

As described in section 1.3, one of the major constraints of epigenetic studies employing pituitary tumours and normal gland is their relatively small size. For the whole genome analyses described in this thesis a range of techniques and their associated end-points were proposed. These end-points included: BeadArray analysis, bisulphite sequencing, pyrosequencing across multiple genes, quantitative reverse-transcript PCR (qRT-PCR), and chromatin immunoprecipitation (ChIP analyses). Of these, the first three requires DNA; qRT-PCR requires mRNA; and ChIP requires chromatins. Therefore, to reduce the impact imposed by sample limitations, I first optimised and validated the technique Whole Genome Amplification of Bisulphite Converted of DNA (BSC-WGA-DNA). This technique is expected to provide more material for three end-point analysis that works on DNA mentioned above.

3.1. Whole genome amplification of bisulphite converted DNA

The literature describes two whole genome amplifications (WGA) methods. The first, primer extension pre-amplification (PEP) which is a PCR-based technique using *Taq* DNA polymerase and random primers to amplify genomic DNA (Zhang *et al.*, 1992). The second method, multiple displacement amplification (MDA), is a non-PCR technique which employs an isothermal genome amplification of ϕ 29 DNA polymerase to amplify genomic DNA (Dean *et al.*, 2002). Of these two methods, since the PEP technique has been used previously by other investigators (Mill *et al.*, 2006) to amplify bisulphite converted DNA (BSC) this method was chosen for preliminary investigation and in the

subsequent studies described in this thesis. However, in this case the WGA amplification technique is a modification of a technique first described by Zhang *et al* (Zhang *et al.*, 1992). For convenience, throughout the rest of this thesis, bisulphite converted DNA either prior to or following WGA are termed WGA- and WGA+, respectively.

WGA was performed as described in M&M section and the result is presented in figure 3.1.

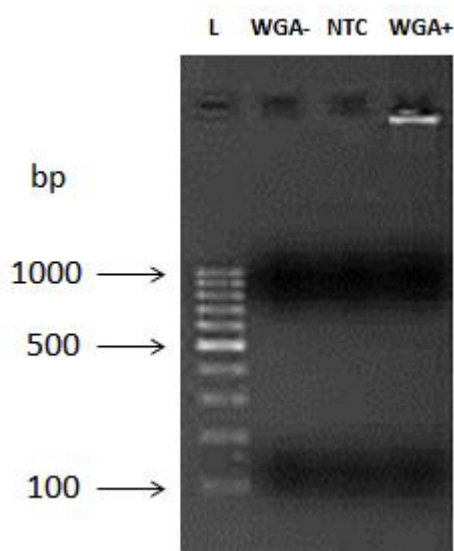
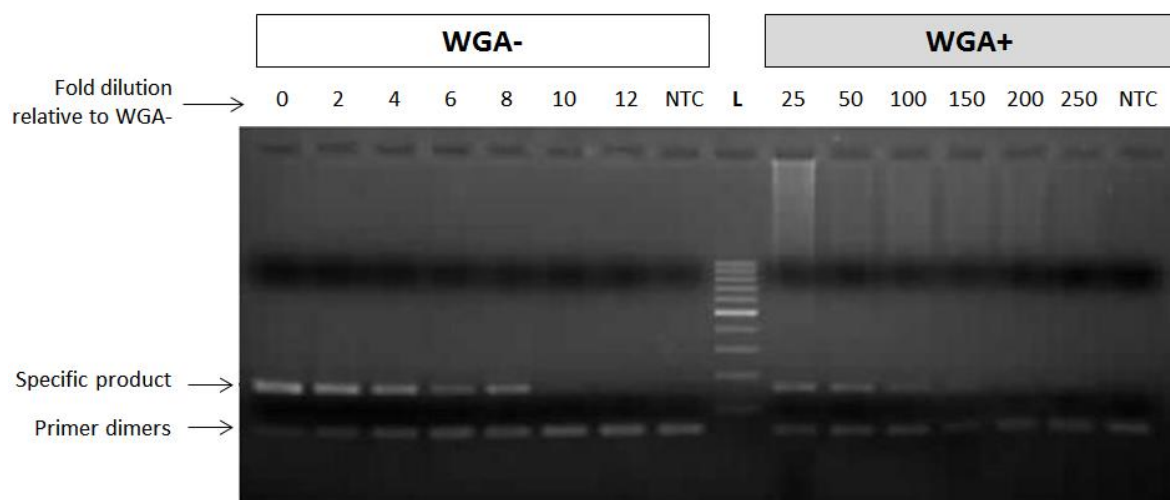


Figure 3.1. Whole genome amplification results in amplification of bisulphite converted DNA

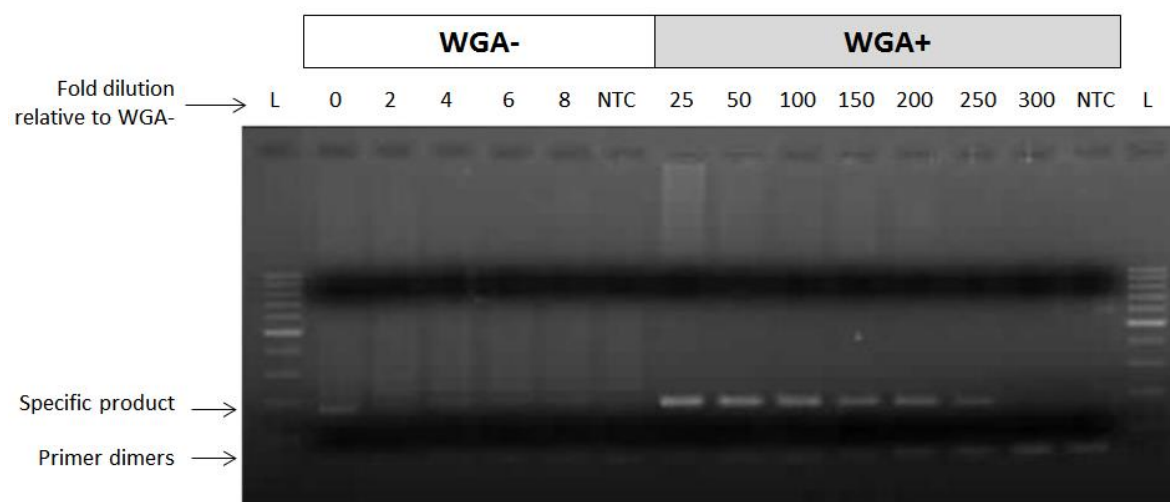
20 ng of BSC DNA derived from a primary pituitary tumour was used as template for WGA reaction. The equal amount of BSC DNA (WGA-) is not visible. The non-template control (NTC) is not visible indicating that there is no contamination during the experiment. The whole genome amplification (WGA+) results in a clearly visible band and slight smear located closely to the loading well. This indicates that BSC DNA was efficiently amplified by WGA technique.

In figure 3.1, the DNA from the non-template control WGA reaction (lanes 3) is not visible, demonstrating that there is no contamination of the PCR reaction. In lane 2, 20 ng of converted, but not amplified DNA (WGA-), is not visible on the gel, however this amount was efficiently amplified through WGA (WGA+) that is evident as an intense, unresolved band, located mainly within the electrophoresis well (lanes 4).

Subsequently, both WGA- and WGA+ were subjected to sequence- and BSC-specific nested PCR amplification and the results are presented in figure 3.2. In these examples different primer pairs, showing different, relative, amplification efficiencies are shown in panel A and B respectively.



(A)



(B)

Figure 3.2. Representative examples of sequence-specific nested PCRs using dilution series of WGA- and WGA+ as templates.

(A) A dilution series from 0 to 12 of 20 ng of BSC DNA derived from a single primary pituitary tumour was used as templates for PCR amplification using *GLRX* gene specific primers (left panel). A specific amplicon is visible in dilutions that ranged from 0 to 8. However, this is invisible at fold dilution of 10 and above. In WGA+ series (right panel), PCR using dilution series of WGA derived from the BSC DNA in the left panel were succeeded at dilution folds from 25 to 100 relative to the amount used for WGA reaction which is equal to the one used in WGA- lane 0 (20 ng).

(B) In the case of *MBD* primers, the fold dilutions of WGA- and WGA+ which showed clearly visible products are 0 and 200, respectively.

The success of nested PCR using not only WGA+ but also WGA- as templates shows some differences in amplification efficiencies. However, a range of successful amplification is apparent in dilutions that ranged from 25 fold to 500 fold of WGA+ relative to WGA-. As example, for the *GLRX* gene specific primers (Fig 3.2A) an amplicon is apparent at 25 and 50 fold dilutions in the lanes denoted as WGA+, whereas in the lanes denoted as WGA- dilutions in excess of 8 fold were not visible on the gel. For amplifications shown in panel B, using in this case primers to the *MBD3* gene were successful at dilutions up to 200 fold in the WGA+ amplified DNA. Conversely, in the WGA- DNA amplification in excess of 0 fold were not detected (figure 3.2 B). In most cases, a 25 dilution fold WGA+ DNA relative to DNA that had been converted, but not amplified, DNA (WGA-) reproducibly provided high quality PCR products for the majority of primers used in this study. On this basis this dilution factor was used for the experiments presented in this thesis. In these cases, therefore, the WGA efficiently increased the limitation on DNA availability by 25 fold .

3.2. Integrity of methylation profile of BSC DNA prior to and following WGA

The integrity of DNA methylation sequence profiles of WGA- relative to WGA+ DNA were assessed by two independent techniques that included Sanger sequencing and Pyrosequencing.

3.2.1. Sanger sequencing

PCR products of a methylated gene (*MBD3*) and non-methylated gene (*GLRX*) derived from WGA- and WGA+ templates were cloned into pGEM-T vector for Sanger sequencing. Result is presented in Figure 3.3. For the *MBD3* gene, all of the CpG sites of three PCR clones derived from WGA- and of four PCR products derived from WGA+

were methylated. For the *GLRX* gene, three CpG sites among a total of 16 CpG sites from WGA- were methylated and a single CpG site in the WGA+ were methylated. Although the number of sequenced clones is somewhat limited, the data shows a consistent methylation pattern between PCR products derived from WGA- and WGA+.

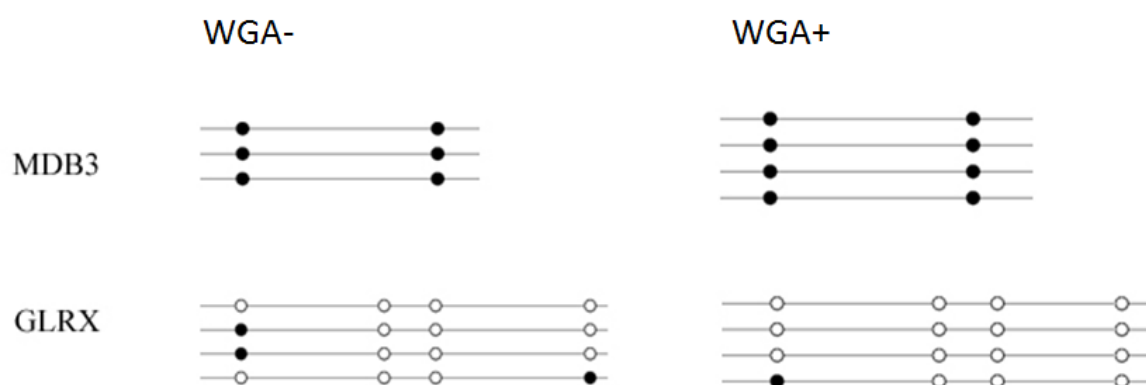


Figure 3.3. Sanger sequencing of *MBD3* and *GLRX*.

PCR products from figure 3.2 were cloned and sequenced to determine their methylation status. Beads on a string depiction of single DNA molecules post T:A cloning are shown in each case strand. Filled circles (beads) are methylated CpGs, un-filled circles are non-methylated CpGs. In the case of *MBD3* gene, all of the CpG sites of three PCR clones derived from WGA- and of four PCR products derived from WGA+ were methylated. In the case of *GLRX* gene, three CpG sites among a total of 16 CpG sites from WGA- were methylated and a single CpG site in the WGA+ were methylated.

3.2.2. Pyrosequencing

Pyrosequencing assays of three genes (*GLRX*, *H19*, and *HAAO*) were designed to interrogate the methylation status of a single CpG per assay for *GLRX* and *HAAO*, and two CpGs per assay for *H19*. Nested PCRs were performed on WGA- and WGA+ of 14 pituitary samples, and PCR products were subsequently subjected to pyrosequencing reactions. Figure 3.4 showed the correlation of 56 pairs of methylation values between WGA- and WGA+ samples. Strong positive correlation was observed ($R^2 = 0.9677$,

$p < 0.0001$). This data confirms that the WGA protocol described in this study is reproducible and reliable.

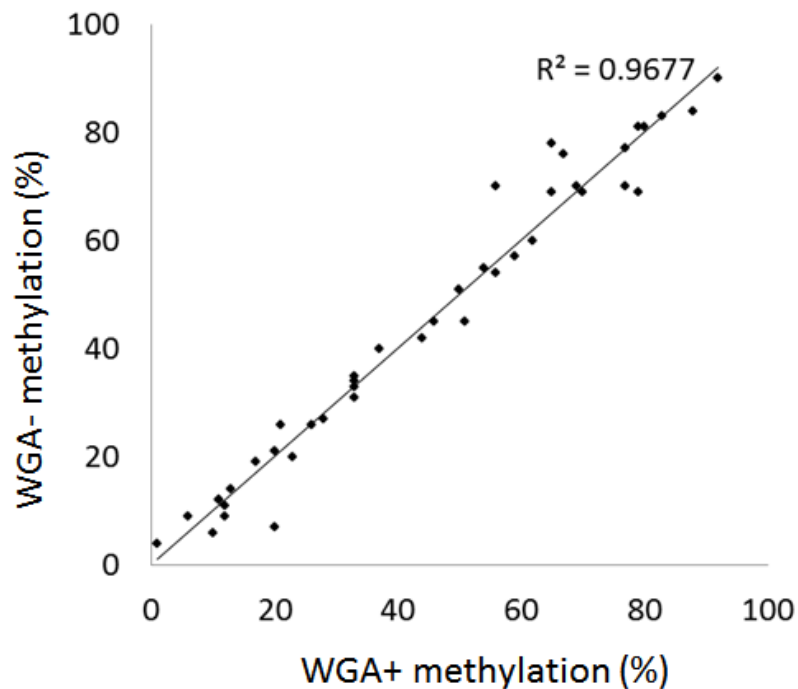


Figure 3.4. Correlation of methylation status produced by pyrosequencing derived from WGA- and WGA+.

Methylation status of a single CpG site of three genes (GLRX, H19, and HAAO) and two CpG sites of H19 were interrogated by pyrosequencing. WGA- and WGA+ of 14 pituitary samples were used for nested PCR. Subsequently, these PCR products were subjected to pyrosequencing reactions that enabled the examination of the correlation of 56 pairs of methylation values between WGA- and WGA+ samples. Strong positive correlation was observed ($R^2 = 0.9677$, $p < 0.0001$).

3.3. Discussion

At the inception of my doctoral studies the technical limitations imposed by tumour size were realised. Furthermore, initial discussions and a review of the literature suggested the importance of performing analyses across and within individual tumours. It was, therefore, considered important to determine epigenetic modifications and expression not only within each of the individual adenoma but also across multiple adenoma subtypes.

Since, in some cases, pituitary adenomas are less than 5mm in diameter, it was necessary to perform whole genome amplification (WGA) to generate sufficient DNA for analysis. However, a caveat associated with this approach is that although it leads to amplification of DNA, either pre- or post- sodium bisulphite conversion, it is not suitable for increasing the relative amount of material available for analyses of histone modification (ChIP) or indeed transcript expression. Therefore, the initial focus of my studies was to employ and optimise the whole genome amplification of sodium bisulphite converted DNA (BSC-WGA) technique to increase the amount of sample DNA. However, it was important to determine, not only the approximate amplification but the integrity of the sequence pre- and post- amplification.

The BSC-WGA protocol described by Mill et al (Mill *et al.*, 2006) was employed to amplify genomic DNA extracted from pituitary tumours. In their study, Mills and colleagues achieved an amplification of some 113 fold following amplification (WGA+) and relative to converted DNA that had not been subject to amplification (WGA). However, in this case the fold amplification was determined on the basis of that obtained with four primer pairs. My own studies employed, in total, more than 50 primer pairs, and across these different primers a range of amplification efficiencies were apparent (data not shown). However, a principle concern was to develop and optimise a technique that was reliable and reproducible. Furthermore, it was important to use conditions where the sequencing of post amplified DNA reflected that observed pre amplification. In this case, and as detailed in the results section, conditions were used that provided a reliable and consistent 25 fold amplification and in this way provided confidence in the technique and in the derived sequencing data.

The integrity of the DNA sequencing profiles following the WGA of BSC converted DNA was determined by two techniques-Sanger sequencing of individual clones

(T:A cloning) and by Pyrosequencing. The first technique, bisulphite sequencing of individual clones showed that methylation profiles were maintained, that is faithfully replicated between WGA+ and WGA- DNA samples. Although I realised that only four Sanger sequencing reactions is a limited number to give a conclusion of unbiased methylation profile prior to and after WGA, I did not perform more Sanger sequencing because of i) similar findings, with respect to reproducibility, were also reported by Mills and colleagues (Mill *et al.*, 2006) and ii) limited source of pituitary DNA. However, I was aware that Sanger sequencing of bisulphite converted DNA has some limitations since the findings are representative of a limited number of clones. In this case the clones analysed are those selected following T:A cloning. Although it is possible to sequence multiple clones this is not only time consuming but also have significant cost implications. To circumvent this limitation Mill *et al* employed methylation-sensitive single nucleotide primer extension (MS-SNuPE). In this case this technique has been reported to provide an accurate assessment of methylation to within approximately, 5% (Kaminsky *et al.*, 2005). In this thesis, and consequent to technical advances I was able to determine methylation using the pyrosequencing technique. In this case the technique and its associated technology is capable of interrogating millions molecules in a single reaction (Colella *et al.*, 2003) and in this way provides a more representative analysis of a poly or multi clonal population of molecules. Again, and in agreement with Mill's results, my data showed significant correlation of the DNA methylation profiles between WGA+ and WGA-. Moreover, the unbiased methylation profile of DNA prior to and after WGA of BSC DNA has been widely reported (Mill and Petronis, 2009; Vaissiere *et al.*, 2009; Mill *et al.*, 2006).

A further anticipated limitation associated with my study was that DNA methylation profiling strategies based on bisulphite treatment have reported the issue of

incomplete conversion of unmethylated C (Grunau *et al.*, 2001; Genereux *et al.*, 2008). To circumvent this issue the conditions for the conversion reaction needs to be harsh, however, this leads to the further concern of large-scale degradation of converted DNA (Grunau *et al.*, 2001). Although techniques are described that are designed to limit degradation of the DNA these protocols requires multiple further manipulations and thus increase the likelihood of introducing errors (Olek *et al.*, 1996). Furthermore, following bisulphite conversion the DNA strands loses their complementarity, since unmethylated C are converted to T and this contributes further to degradation issues. For these reasons commercially available BSC conversion protocols recommend that converted DNA to be used within one week of conversion. However, a key advantage of WGA that we and others have noted (Mill *et al.*, 2006) is that post conversion the DNA again become double stranded and as such not only quantitatively increases the amount of BSC DNA, but also largely improves the stability and storage time of valuable samples.

In conclusion, WGA technique has been successfully adapted toward analysis of sodium bisulphite converted DNA from small samples of tumour material. This technique significantly increased and improved the sodium bisulphite conversion of DNA in both quantitative (typically ~ 25 fold) and qualitative (extended storage) terms. Sequence specific PCR were typically successful using a 25 fold amplification of WGA+ relative to WGA- and the DNA methylation profiles derived from WGA- and WGA+ do not appear to show any bias. This simple and inexpensive technique is recommended for DNA methylation profiling studies where only limited sample are available.

Chapter 4: Genome-wide DNA methylation profiling of sporadic pituitary tumours

The optimisation of the WGA technique provides a reliable and reproducible tool to partially overcome the limitation of sample shortage. Therefore, this facilitated genome-wide DNA methylation analysis of promoter-associated CpG islands and characterisation of their consequent effects on gene expression. To determine the genome-wide subtype-specific methylation profiles of sporadic pituitary adenomas several factors were taken into account and that included, sample throughput, epigenome coverage, sample requirements, and costs. A decision between absolute adenoma number on the array and the individual adenoma subtypes investigated. A decision was made to use the recently commercialised Illumina Infinium 27K Assay. This assay is able to determine the methylation status of 27,578 CpG sites that span 14,496 genes. However, each single array is capable of interrogating 12 samples at most and with considerable cost implications. A further decision was that three samples of each adenoma subtype together with, as control, three post-mortem normal pituitaries would be investigated.

4.1. Methylation profiling of sporadic pituitary adenoma

As described in the preceding paragraph, the discovery cohort comprised adenomas that represented each of the major pituitary adenoma subtypes, non-functioning (NF), somatotrophinoma (GH), prolactinoma (PR) and corticotrophinoma (CT). Quantitative differential methylation was assessed relative to post-mortem normal pituitaries (NP).

Before determining differential methylation, 1,092 CpGs (4%) in chromosomes X and Y were excluded because my pituitary samples were sex-independent. Next, 1,090 low-quality probes (4%), which showed either detection probability >0.05 (internal quality

control) or null (missing) β value, were removed. In addition, 14,729 probes (53%) where β values were either ≤ 0.2 or ≥ 0.8 across the 15 samples were removed to reduce the number of non-variable loci as we and others have previously described (Byun *et al.*, 2009; Fryer *et al.*, 2011). Through application of these criteria, a total of 10,667 probes (39%) spanning 7,956 genes were included in the analysis. Compared to our lab's previous publication (Fryer *et al.*, 2011), 39% is the better percentage. Within this probe set, 5,287 CpG were within and 5,280 were outside of CpG islands. A Kolmogorov–Smirnov test was used by our collaborator (Dr Richard Emes) to compare the distribution of β values between samples using the NIMBL Software (Wessely and Emes, 2012). A single NP sample was identified as having a significantly different ($P < 0.05$) distribution to the other samples and was excluded from further analysis (Appendix 7). To characterise the differential methylation profiles of the adenoma subtypes, we first performed a class comparison between each of the adenoma subtypes and relative to that apparent in the NP. For the comparison, we initially filtered for a minimum increase of $\Delta\beta$ of 0.4 relative to NP at single gene-specific CpG sites in at least two of three subtype-specific adenomas. This analysis identified 326 genes in the NF adenomas, 49 in PRL, 97 in GH and 4 in CT that were hypermethylated relative to NP supplemental data (Appendix 5). In the majority of cases and irrespective of subtype, the CpGs were within CpG islands. These findings were encouraging, in that they show genes that were methylated in common and also that were discrete across and with individual adenoma subtypes respectively. However, it was important that I was able to validate the BeadArray findings using an independent methodology, and this is described below.

4.2. Technical validation of Bead array by pyrosequencing

To validate the BeadArray-generated data, I performed pyrosequencing analysis of six CpGs in six gene-associated CpG islands from the 12 samples (discovery cohort) that had been interrogated on the array. Of these seven CpGs, for technical validation purpose, the pyrosequencing assays of four sites (cg08047457, cg20289949, cg06197492, and cg24739326) were designed directly to target these CpGs which had been interrogated on the BeadArray. However, the other two assays were not designed for technical validation purpose. The reason is that in the later part of my study where the aim is to search for abnormal methylated genes in sporadic pituitary tumours, pyrosequencing assays were preferably designed to target 5 to 11 CpG sites (Appendix 3.2) located immediately before and partly overlap with the first exon of the corresponding genes because this region is usually associated with transcription regulation mechanisms. Because of this strategy of pyrosequencing assay designing, the majority of the CpGs interrogated by BeadArray were not included in pyrosequencing assays. Only two pyrosequencing assays of *COL1A2* and *KCNE3* includes CpGs on the array (cg18511007 and cg02595219, respectively) and therefore these data points were collected and included in this analysis. The analysis (Figure 4.1) for the quantitative measurement of CpG methylation showed the two techniques to be significantly correlated (Spearman's r , 0.89; $P \leq 0.0001$). Importantly, the correlation involves samples across the range of β values and hence is unlikely to be an artefact due to dominance of extreme values (Roessler *et al.*, 2012). In addition, inspection of the 20 imprinted genes interrogated on the array showed a mean β value of 0.51 ± 0.1 across a total of 40 CpG sites. These β values are consistent with monoallelic methylation apparent in most normal tissues including pituitaries and providing further confidence in the methodological approach.

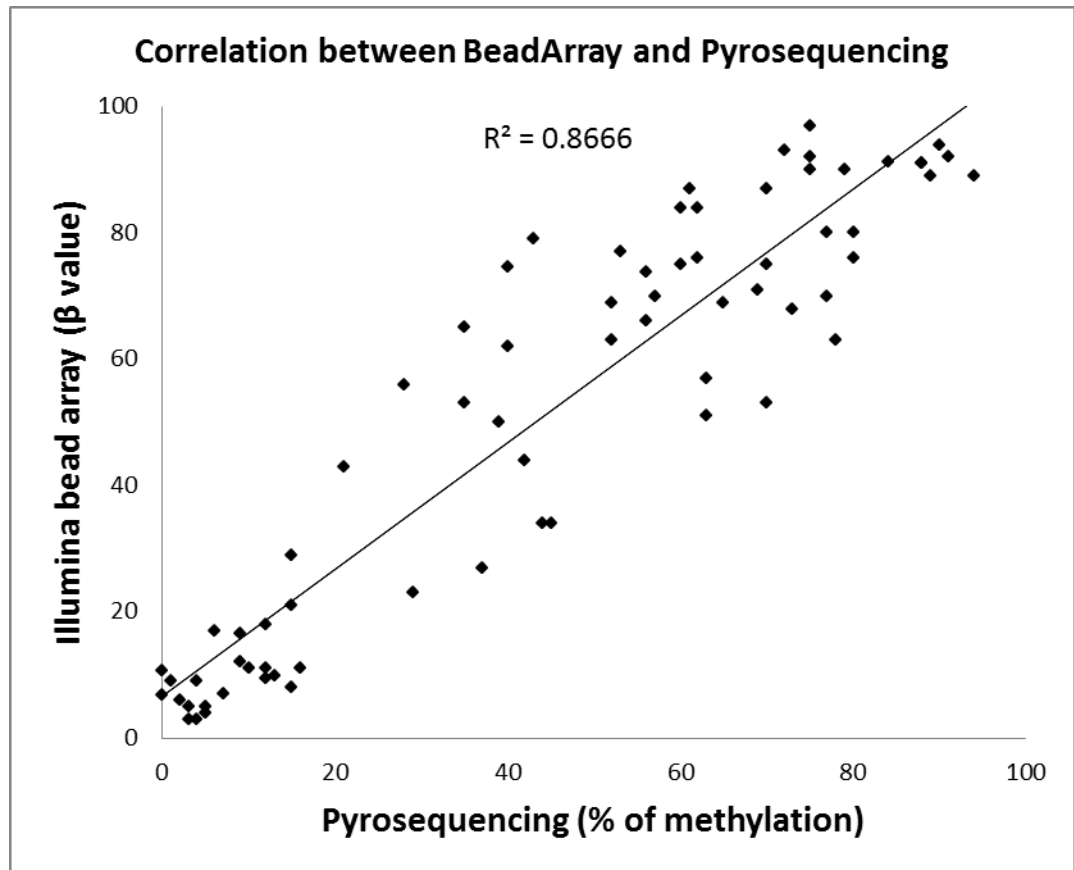


Figure 4 1. Correlation between BeadArray-derived β values (expressed as a percentage) and pyrosequencing across a range on β values from 12 adenomas for four CpGs

The beta values (derived from BeadArray) and the methylation values (derived from pyrosequencing) of 6 CpGs (cg08047457, cg20289949, cg06197492, cg24739326, cg18511007, and cg02595219) of 12 pituitary tumours were compared. Correlation of 72 data points of data shows good correlation (Spearman's r , 0.89; $P \leq 0.0001$) between the two techniques that are BeadArray and Pyrosequencing. The aim of this experiment is to technically validate the reliability of BeadArray assay, therefore, data points above 0.8 and below 0.2 were not excluded as described in section 4.1.

4.3. CpG island methylation profiles in primary adenoma subtypes

To determine the relationship between methylation at array-interrogated CpG sites and methylation across the gene-associated CpG islands, pyrosequencing analysis was used that encompassed, in each case, between 5 and 11 CpGs (Appendix 3.2). As described in section 4.2, the CpGs interrogated by BeadArray were rarely included in pyrosequencing assay, except for the two cases of *KCNE3* and *COL1A2*. Initially, 22 genes were selected on the basis that they each showed a $\Delta\beta$ of ≥ 0.4 (in at least two of the three NF adenomas) relative to NP and where the array-interrogated CpG are within promoter-associated CpG islands (Table 4.1). The analysis was initially performed in the discovery cohort of adenomas and relative to four NP. Mean methylation across these CpGs, as determined by pyrosequencing, was able to confirm the BeadArray methylation status for only 12 of the 22 genes (55%). However, it is interesting to note that where concordance between the techniques is observed (12 of 22 genes), it was for genes that were represented by more than one island-associated CpG on the array, and where the $\Delta\beta$ for the second CpG also showed increase.

Gene	BeadArray		Pyrosequencing
	Criteria		Number of NF tumours showing increase in mean methylation of CpG islands relative to NPs
<i>EML2</i>	1st CpG ≥ 0.4	2 nd CpG ≥ 0.25	2
<i>RHOD</i>			2
<i>HOXB1</i>			1
<i>FLJ46380</i>			2
<i>MT1G</i>			2
<i>HAAO</i>			2
<i>KIAA1822</i>			3
<i>KCNE3*</i>			3
<i>TFAP2E</i>			2
<i>BCL9L</i>			1
<i>PGLYRP1</i>			2
<i>COL1A2*</i>			2
<i>ST6GALNAC6</i>			0
<i>UBTD1</i>			0
<i>KIAA0676</i>			0
<i>PDLIM4</i>			0
<i>CXCL14</i>		2 nd CpG < 0.25	0
<i>FEM1C</i>			0
<i>MAL</i>			0
<i>RAMP1</i>			0
<i>D4ST1</i>			0
<i>SAMD11</i>			0

Table 4.1. Identity of 22 randomly selected genes from the BeadArray dataset.

22 of the genes were selected on the basis that they had at least one CpG within the gene (in at least two of three specimens of an adenoma subtype) with a β value of ≥ 0.4 relative to the values seen in NP. Of these 22 genes 16 also harboured a further (second) CpG where the β value was ≥ 0.25 . For the 6 genes selected on the basis that they had a single CpG within the gene (in at least two of three specimens of an adenoma subtype) with a β value of ≥ 0.4 the second CpG was ≤ 0.25 and relative to the values seen in NP. Genes where pyrosequencing validated their methylation status of the promoter-associated CpG islands are shown in bold. Genes indicated by a (*) possess a BeadArray CpG which is also interrogated by Pyrosequencing.

Pyrosequencing was performed on the remained DNA of the 3 NF and 2 NPs of the discovery cohort returned from Illumina.

For the 22 genes selected for pyrosequence analysis, on the basis of a single CpG site, 16 genes also showed increase in a second CpG site ($\beta \geq 0.25$), in at least two of the three NF adenomas that had also been interrogated on the array (Table 4.1). Of these 16 array-identified genes, pyrosequence analysis confirmed increase in mean methylation (across 5 to 11 CpG sites) for 12 of 16 (75%) genes in the discovery cohort of NF adenomas. It is worth to note that in cases of two genes (*KCNE3* and *COL1A2*) where a CpG interrogated by BeadArray were also included in pyrosequencing, the β -value of this CpG determined by BeadArray is not only strongly correlated with its methylation level determined by pyrosequencing (Figure 4.1), but also reflects the methylation status of the surrounding CpG regions. However, because of the limited number of these cases, it is risky to give a conclusion. The aim of my phd study (Section 1.4) was not designed for technical validation of BeadArray assay. It was finding novel genes related to tumourigenesis of sporadic pituitary tumours. Therefore, there should be another study to deal with the question of whether methylation of a single CpG correctly reflect methylation of it's surrounding CpGs region.

On the basis of the finding described earlier, we modified our filter criteria to include genes where one of the two CpGs showed a $\Delta\beta$ of ≥ 0.4 and a second CpG a $\Delta\beta$ of ≥ 0.25 and where these criteria were fulfilled in at least two of the three adenomas (within a subtype) interrogated on the array. The revised, more stringent, filtering criteria identified 40 genes in NF adenomas, 21 in GH, 6 in PRL and 2 in CT that showed increase in methylation-relative to NP (Table 4.2). Table 4.2 also shows genes that are methylated in common across adenoma subtypes and those that segregate with one or more pituitary adenoma subtypes. In addition, and for two of the adenoma subtypes, NF and GH, where these criteria were met in multiple genes, we generated a heatmap displaying β values across individual adenomas within these subtypes (Figure 4.2).

Gene	GH	NF	PRL	CUSH
<i>SEPT9</i>				
<i>ACTA1</i>				
<i>ADRA1A</i>				
<i>ALX4</i>				
<u>BCL9L</u>				
<i>BIK</i>				
<i>BMP8A</i>				
<i>C1QTNF5</i>				
<i>C6orf150</i>				
<i>CHST8</i>				
<i>CNFN</i>				
<u>COL1A2</u>				
<i>CRIP1</i>				
<i>DLX5</i>				
<i>EFEMP1</i>				
<i>EFS</i>				
<i>ELN</i>				
<u>EML2</u>				
<i>ENTPD2</i>				
<i>ERBB2</i>				
<i>FLJ32569</i>				
<i>FLJ32569</i>				
<u>FLJ46380</u>				
<i>FLJ90166</i>				
<u>HAAO</u>				
<i>HAS1</i>				
<i>HDAC11</i>				
<u>HOXB1</u>				
<u>KCNE3</u>				
<i>KCNQ1</i>				
<i>KIAA0676</i>				
<u>KIAA1822</u>				
<i>LDHC</i>				
<u>MT1G</u>				
<i>NPPB</i>				

<i>PDLIM4</i>				
<u>PGLYRP1</u>				
<i>PON3</i>				
<i>RAB34</i>				
<i>RAC2</i>				
<i>RASSF1</i>				
<u>RHOD</u>				
<i>SECTM1</i>				
<i>SIPA1</i>				
<i>SLC5A1</i>				
<i>SLITL2</i>				
<i>SOCS1</i>				
<i>SOCS2</i>				
<i>ST6GALNAC6</i>				
<i>TCF15</i>				
<u>TFAP2E</u>				
<i>TNFRSF10D</i>				
<i>UBTD1</i>				
<i>VILL</i>				
<i>WFIKKN2</i>				
Total	21	40	6	2

Table 4.2. Genes identified on the BeadArray as hypermethylated.

Genes were identified as hypermethylated on the basis of one of two CpGs (in two of three adenomas within a subtype) showing a β value of ≥ 0.4 and a second CpG showed a $\beta \geq 0.25$ relative to the mean of the NP. The table shows the genes fulfilling these criteria in each of adenoma subtypes and where the CpG is within a CpG island. Genes shown in bold and underlined were subject to the additional analyses described in Figure. 4.3 and Table 4.3

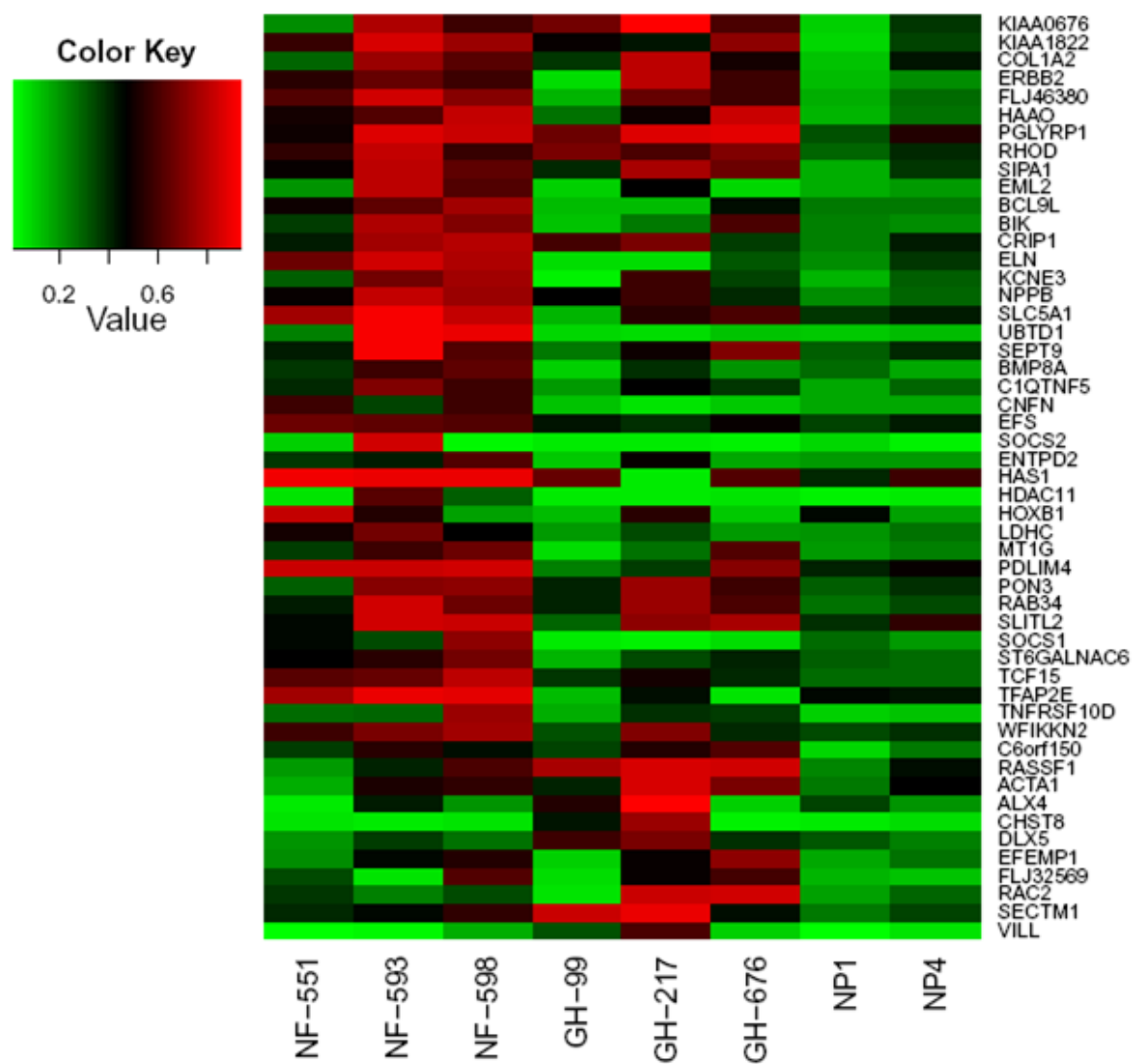


Figure 4.2. Heatmap of genes in individual NF and GH adenomas that fulfilled the criteria of elevated β in two gene-associated CpGs relative to the NP.

For each gene that satisfied the criteria, the mean of the two β values is used and displayed as a graded colour scale that is shown in the figure.

4.4. Gene specific methylation frequencies in non-functioning adenomas

Through constraints imposed by methodologies (sodium bisulphite conversion, pyrosequencing and transcript expression analyses), the availability of some adenoma subtypes (number and amount of tumour tissue) with the exception of NF adenomas was limited. I, therefore, focused this part of the study primarily, but not exclusively, on the NF adenoma cohort. Through analysis of the 12 genes identified by BeadArray and confirmed by pyrosequencing in the discovery cohort, I determine the frequency of aberrant methylation across, in this case, an independent cohort of 13 NF adenomas relative to four NP. Figure 4.3 shows the analysis where each gene in the NP, with the exception of *HOXB1*, shows low but variable levels of methylation. I set a stringent cut-off criterion for increased methylation in individual adenomas of ≥ 4 S.D. higher than the S.D. of the mean (for each of the genes) that was apparent in the four NP. The proportion of adenomas showing increase, above each of the gene-specific thresholds, is shown in Figure 4.3 and is summarised in Table 4.3. Within the NF cohort, the number of adenomas that show gene-specific increase in methylation is variable, where the CpG island methylation of the *KIAA1822* (*HHLPL1*) and *TFAP2E* is found in 12 of 13 NF adenomas whereas the *COL1A2* CpG island shows increased methylation in only two of 13 adenomas.

#	Gene	Hypermethylated tumours	Normal methylated tumours
1	<i>BCL9L</i>	4	9
2	<i>COL1A2</i>	2	11
3	<i>EML2</i>	7	6
4	<i>FLJ46380</i>	9	4
5	<i>HAAO</i>	11	2
6	<i>HOXB1</i>	7	6
7	<i>KCNE3</i>	10	3
8	<i>KIAA1822</i>	12	1
9	<i>MT1G</i>	11	2
10	<i>PGLYRP1</i>	7	6
11	<i>RHOD</i>	8	5
12	<i>TFAP2E</i>	12	1

Table 4.3. Frequencies of gene-specific methylation in an independent cohort of NF adenomas.

Genes were first identified in the discovery cohort and using the criteria described in the text and in Figures 4.2 and 4.3

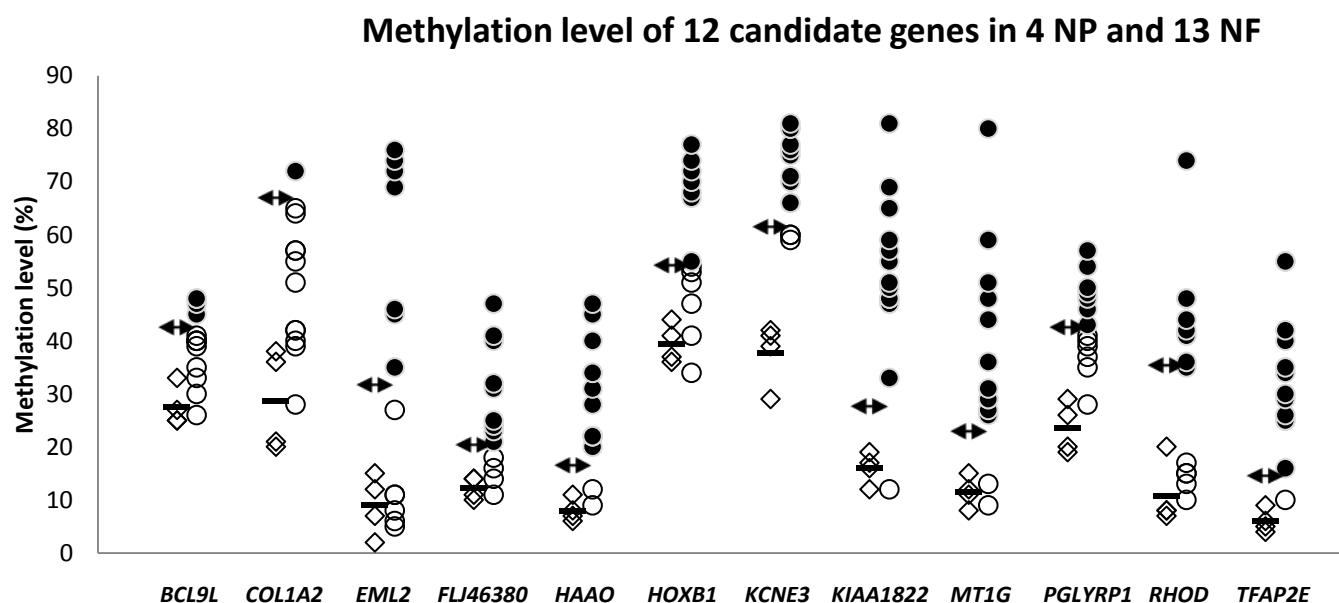


Figure 4.3. Pyrosequence analysis of promoter-associated CpG island methylation of genes first identified as hypermethylated by BeadArray analysis.

The figure shows methylation in an independent cohort of 13 NF adenomas relative to the mean methylation for each gene examined in the four NP. For each gene, in the NP, the mean percentage methylation and four times the S.D. of the mean are shown as horizontal bars, either without or with arrows respectively. Only adenomas equal to or greater than this upper threshold were deemed hypermethylated and are shown as filled circles. Adenomas in which methylation is not increased relative to NP are shown as unfilled circles. The NP are shown as diamonds to the left of the adenomas (circles) in each case.

4.5. Subtype specific gene methylation frequencies

Of the 12 genes subject to validation in the NF adenomas, six of these genes also fulfilled the $\Delta\beta$ criterion in two or more of the other adenoma subtypes (Table 4.2). To gain insight into the subtype specificity of hypermethylation, I performed pyrosequence analysis for three of these genes where, in addition to the NF adenomas, I included each of the other pituitary adenoma subtypes (Figure 4.4). For *EML2*, and confirming the BeadArray finding, methylation is confined to NF adenomas. Cross subtype-specific methylation is apparent for the *RHOD* and *KIAA1822* genes where the array (Table 4.2) and pyrosequencing shows increased methylation in a proportion of the NF- and GH-secreting adenomas and in the NF-, GH- and PRL-secreting adenomas respectively (Figure 4.4).

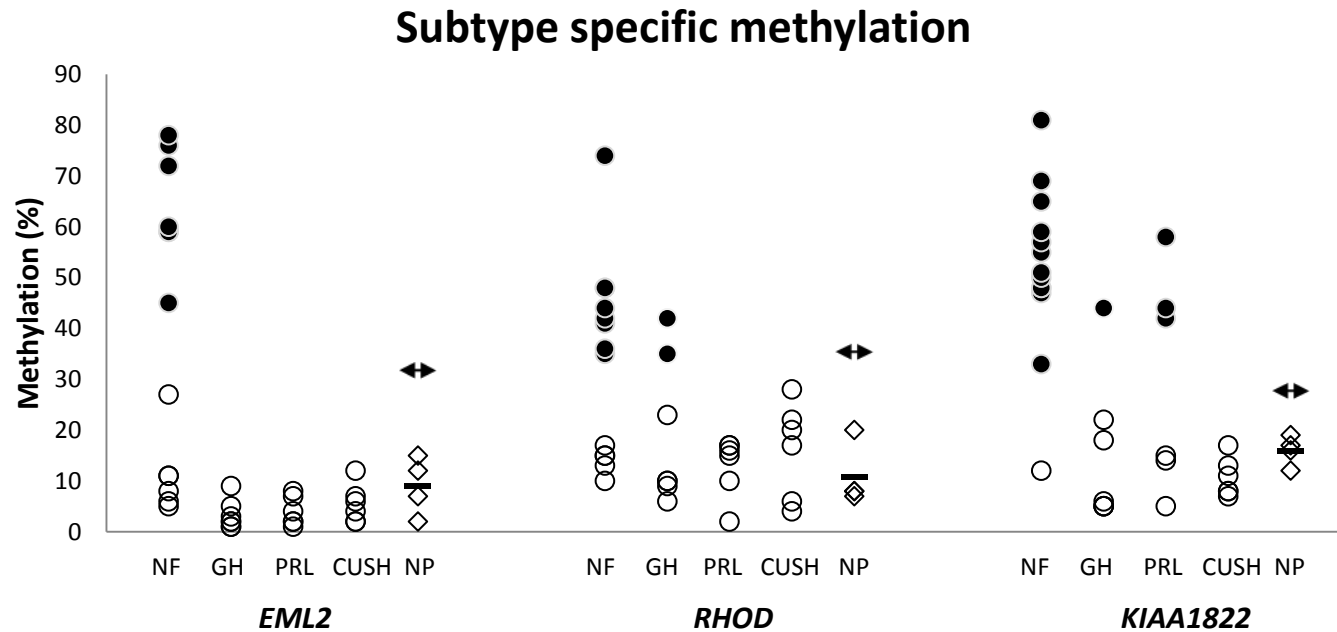


Figure 4.4. Adenoma subtype specificity of hypermethylation for three genes in which pyrosequencing confirmed their methylation status in NF adenomas is shown in Figure 4.3.

For subtype specificity, in addition to the 13 NF adenomas, the methylation status of each gene was also determined in an investigation cohort comprising, seven GH, six PRL and six CT (CUSH) adenomas. Symbols and cut-off criteria are as described in Figure 4.3. The NP are shown as diamonds on the right side of each of the gene investigated.

4.6. Methylation associated gene silencing

For the 12 genes where pyrosequencing confirmed methylation status, I used quantitative RT-PCR (RT-qPCR) to determine associations between CpG island methylation and gene silencing. In these cases, loss or significantly reduced transcript expression was assigned on the basis of the stringent criterion described for gene-specific methylation. However, in these cases, it was ≤ 4 S.D of the NP. In the 13 NF adenomas, an inverse relationship between increased methylation and reciprocal significant decrease in transcript expression is apparent for three of the 12 genes examined. Figure 4.5 shows that relative to NP, all the methylated tumours (filled circles) show significantly reduced transcript expression. Conversely, in adenomas that were not hypermethylated (unfilled circles), expression levels are similar to or in some cases higher than that apparent in the NP. A single exception to these findings is apparent for the *HOXB1* gene, where a single adenoma that was not methylated also shows reduced transcript expression. In addition, as multiple studies using candidate gene approaches have shown frequent methylation-associated loss of *CDKN2A* (Woloschak *et al.*, 1997; Simpson *et al.*, 1999; Vandeva *et al.*, 2010), we included RT-qPCR analysis of this transcript. These studies show significantly reduced expression of this transcript in 12 of 13 adenomas relative to NP (Appendix 8) further confirming and reinforcing the robustness of the approach.

Associations between Methylation and Gene Expression

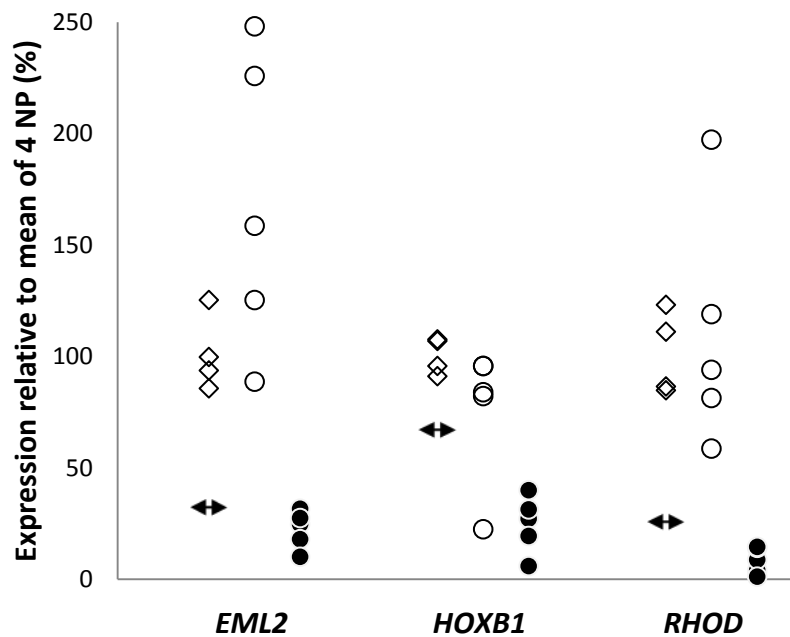


Figure 4.5. Quantitative RT-PCR analysis for three of 12 gene transcripts that show an inverse relationship between hypermethylation and decrease in transcript expression.

For each gene, their methylation status in NF adenomas was first validated in the discovery cohort (Table 4.2) and in the independent cohort shown in Figure. 4.3. Expression is shown relative to four NP and the same cut-off criteria of four S.D., but in this case lower than the mean level shown in NP (horizontal bars with arrow heads). Filled and unfilled circles represent adenomas that are hypermethylated and those that were not for individual NF adenoma respectively. The NP are shown as diamonds on the left side of each the gene investigated. Relative to the housekeeping gene *GAPDH*, mean expression levels of *EML2*, *HOXB1*, and *RHOD* in 4 NPs, as estimated by absolute Ct values in RT-qPCR, are approximately 1%, 0.01%, and 1%, respectively.

4.7. Gene ontology analysis of hypermethylated genes

Gene ontology analysis of the hypermethylated genes was performed to determine the biological processes where these genes are known to contribute together with their molecular function and finally their associated Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. These analyses are shown in Appendix 6. The aberrantly methylated, promoter-associated CpG islands were distributed across various categories of biological processes, molecular function and pathways. However, multiple genes associated with particular biological processes showed frequent methylation in the analysis. In this context, we noted that 11 genes in intracellular signalling cascades (Appendix 5.1) were hypermethylated in pituitary adenomas. Interestingly, among these genes, the majority (*SOCS2*, *RAC2*, *ERBB2*, *RASSF1*, *SOCS1* and *COL1A2*) have also been shown to exhibit tumour suppressor properties in multiple other tumour types. *RASSF1* was shown to be frequently inactivated by hypermethylation of its promoter region in a number of tumour types including bladder cancer, breast cancer, cervical cancer, and pituitary adenomas (Mitra *et al.*, 2012; Jiang *et al.*, 2012; Ha *et al.*, 2012; Gil *et al.*, 2012; Qian *et al.*, 2005). Higher expression level of *SOCS1* is associated with earlier tumour stage and better clinical outcome in human breast cancer (Sasi *et al.*, 2010). *COL1A2* frequently silenced by promoter hypermethylation in melanoma (Bonazzi *et al.*, 2011). *RAC2* is critical to the initiation of acute myeloid leukemia in a retroviral expression model of *MLLAF9* leukemogenesis (Mizukawa *et al.*, 2011). *ERBB2* methylation status is strongly associated with an epithelial phenotype in NSCLC cell lines, surgically resected tumours, and formalin-fixed biopsies from patients with non-small-cell lung cancer who went on to fail frontline chemotherapy (Walter *et al.*, 2012). KEGG pathway analysis shows multiple genes associated in various types of cancer (bladder cancer, non-small cell lung cancer,

pancreatic cancer, colorectal cancer, endometrial cancer, prostate cancer) or tumorigenesis-related pathways (apoptosis, VEGF signalling).

4.8. Discussion

The study of the epigenome in normal and disease states, and in particular as it relates to tumourigenesis, is an area of growing interest and research focus. In this thesis, the role of changes in DNA methylation at CpG dinucleotides has been analysed in each of the major pituitary adenoma subtypes. Through use of an Infinium BeadArray approach, I performed a genome-wide analysis, thus providing a quantitative measure of DNA methylation at CpG sites across the genome. The analysis revealed aberrant methylation in each of the adenoma subtypes. Furthermore, it showed that hypermethylation, at CpG sites within gene-associated CpG islands, is more frequent compared with CpG sites outside of CpG islands and similar findings in other tumour types have been reported elsewhere (Kim *et al.*, 2011a). The epigenetic changes were more frequent in NF adenomas than other pituitary adenoma subtypes; however, I recognise that the absolute numbers of sites identified as hypermethylated is dependent on the cut-off criteria employed.

As in our previous studies and those of other groups, the technical validation of the BeadArray approach shows it to be a robust approach for the quantitative determination of methylation (Fryer *et al.*, 2011; Lima *et al.*, 2011). Furthermore, and although not reported by others, I found, in NP, that the imprinted genes show β values that reflect silencing of one of two alleles and provides further confidence in the derived data.

For the majority of genes interrogated by the 27K BeadArray technology, most are represented by one or two CpG sites on the array (Irizarry *et al.*, 2008; Wilop *et al.*, 2011). In these cases, although the number of CpG sites interrogated, as representations of a CpG

islands, are few in number, the *a priori* of the technology is that CpG sites in close proximity are often co-methylated (Ogoshi *et al.*, 2011; Irizarry *et al.*, 2008). However, my validation studies on the basis of an elevated β value at a single CpG site failed to show a robust correlation. In these cases and by pyrosequence analysis, I had determined mean methylation across 5–11 CpGs within the promoter-associated CpG islands. Using modified and more stringent criteria, which considered two promoter-associated CpGs on the BeadArray, I found significant concordance between the techniques. In this case, 12 of 16 (75%) genes were validated by pyrosequencing. Other studies, where they report validation of BeadArray data, frequently use techniques such as methylation-sensitive PCR (Wilop *et al.*, 2011; Yoon *et al.*, 2010). A limitation of this technique is that they determine methylation at a limited number of CpG sites. Equally, this approach is reliant upon an amplification step and may not necessarily accurately reflect the methylation status of the CpG islands that is representative of the majority of cells within the specimen (Zeller *et al.*, 2012; Irizarry *et al.*, 2008; Wilop *et al.*, 2011).

The gene ontology analysis, on the basis of the more stringent cut-off criteria, shows the genes to be distributed across various categories of biological processes, molecular function and pathways. Although no biological processes are specifically enriched, I found that multiple genes involved in intracellular signalling pathways are hypermethylated relative to NP. Interestingly, many of these genes are silenced in association with CpG methylation in other tumour types (Bonazzi *et al.*, 2011; Mizukawa *et al.*, 2011; Mottok *et al.*, 2007; Qian *et al.*, 2005; Sasi *et al.*, 2010; Walter *et al.*, 2012) and interestingly the *RASSF1* gene has been identified in earlier studies using a candidate gene approach of sporadic pituitary adenomas (Qian *et al.*, 2005).

Although the BeadArray allowed me to examine each of the major pituitary adenoma subtypes in a single experiment, which is on the same array, I were aware of the

constraints and limitations of any conclusions reached through examining a limited number of each of the adenoma subtypes. Equally the amount and absolute number of each adenoma subtype were also a limitation for all the analyses I wished to undertake. However, through focusing my studies, post-BeadArray analysis, principally but not exclusively on NF adenomas, I gained further insight into the aspects of epigenetic change. In this case, in an independent cohort that comprised 13 NF adenomas, and for 12 genes (identified in the discovery cohort), I determined methylation frequencies relative to NP. Again using stringent criteria (four times the S.D. apparent in NP), I determined the frequencies of differential methylation across the cohorts for each of these genes. The genes show variable frequencies of methylation and have not been previously described in pituitary adenomas. Indeed, for many of the genes identified in our discovery cohort, as example, *SOCS1*, *SEPT9*, *PDLIM4*, *TFAP2E*, *MT1G*, *HAAO*, *TFAP2E*, *CRIP1* and *COL1A2* methylation-mediated gene silencing, together with known or putative tumour suppressor characteristics has been reported in other tumour types (Bonazzi *et al.*, 2011; Mizukawa *et al.*, 2011; Mottok *et al.*, 2007; Qian *et al.*, 2005; Sasi *et al.*, 2010; Walter *et al.*, 2012).

Although the absolute number and amount of tumour tissue, for the other adenoma subtypes, was limiting, I considered it important to validate the subtype specificity of aberrant methylation. To that end, I examined hypermethylation of CpG islands that were methylated in common and those where this change is apparent in two or more subtypes as determined by BeadArray analysis. The analysis shows, by an independent technique, namely pyrosequencing, that the identified changes and differences in hypermethylation were indeed subtype specific.

In this first report of unbiased genome-wide changes in methylation, I focused my studies on increase in methylation of CpG islands associated with gene promoters. In these

cases, these regions are considered the most relevant for expression of the corresponding gene (Irizarry *et al.*, 2008; Moran *et al.*, 2012). In the 13 NF adenomas and across the 12 genes studied, only three showed an association between hypermethylation and reduced transcript expression. Similar findings have been reported by others and specifically where associations between increased BeadArray-identified CpG methylation and loss or reduced transcript expression has been investigated (Irizarry *et al.*, 2008; Kim *et al.*, 2011a; Wilop *et al.*, 2011; Zeller *et al.*, 2012). Multiple factors might account for this lack of association, including density of methylation across the CpG islands, histone modification or the milieu of transcription factors within a cell or cell type (Irizarry *et al.*, 2008; Wilop *et al.*, 2011). These findings have led some authors to conclude that mRNA expression is more likely to be down regulated when the promoter region is hyper-methylated even though it is not statistically significant (Kim *et al.*, 2011a). Therefore, and by extension, this likely reflects the view put forward by other investigators that many epigenetic and genetic alterations in cancer are ‘passenger’ as opposed to ‘driver’ events in the evolution and progression of a tumour (Kim *et al.*, 2011a; Zeller *et al.*, 2012).

In conclusion, this chapter provides a first, unbiased survey of the pituitary tumour epigenome across different adenoma subtype. Further studies will be needed to corroborate identification of these CpG islands, characterised on the basis of their hypermethylation and relative to that seen in normal pituitaries. Because of the limitation of the number of pituitary samples interrogated by the Infinium 27K array, this study does not provide a biomarker that is good enough for clinical diagnostic applications. However, the evolution of techniques, such as 27K to 450K array, coupled with larger number of samples will make these types of studies useful for identification of biomarker that perhaps predict or characterise tumours likely to show aggressive or recurrent growth characteristics. In addition, the functional characterisation of down-regulated genes will provide new

understanding of tumour aetiology and biology and perhaps identify novel genes or pathways that provide us with new therapeutic targets or options.

Chapter 5: The *EFEMP1* gene: a frequent target for epigenetic silencing in multiple human pituitary adenoma subtypes

As an extension to my genome-wide investigation I sought to clarify the functional relationship between epigenetic changes and gene expression. For these studies I focused attention on the EGF containing fibulin-like extracellular matrix protein 1 gene (*EFEMP1*). My choice of *EFEMP1*, also known as fibulin-3, was on the basis that the BeadArray analysis had revealed subtype-specific differential methylation that was confined to the GH secreting adenomas. I, therefore, considered it may represent an important subtype-specific aberration. In other tumour types *EFEMP1* expression patterns are associated with either growth-promotion or growth-inhibition in a cell- or tissue- type context (de Vega *et al.*, 2009; Gallagher *et al.*, 2005; Obaya *et al.*, 2012). In those cases where cell- and or tumour-type are associated with loss or reduced *EFEMP1* transcript expression it is frequently, but not invariably, associated with CpG island methylation (Hu *et al.*, 2012; Kim *et al.*, 2011b). Moreover, multiple studies show that *EFEMP1* impacts on multiple distinct cellular pathways and these properties provided several functional endpoints for investigation in a pituitary context (de Vega *et al.*, 2009; Sadr-Nabavi *et al.*, 2009; Seeliger *et al.*, 2009). Furthermore, and on the basis of my preliminary studies of pituitary tumour cell lines these revealed that they do not express *EFEMP1* thus providing useful models to gain functional insight into potential mechanisms associated with gene silencing. In these cases I was able to determine the effects of epidrug as mediators of re-expression, and also to determine the functional sequelae of heterologous *EFEMP1* expression within these cells.

5.1. Expression of *EFEMP1* in pituitary adenomas

I first determined expression of *EFEMP1* in an investigation cohort of adenomas comprising each of the major adenoma subtypes. Figure 5.1 shows that by RT-qPCR analysis significantly reduced expression of *EFEMP1* (by the criteria described in the material and methods), in 25 of 33 adenoma (~75%) and relative to expression in post-mortem normal pituitaries. Although the number of subtype-specific adenomas available for each of the analyses is relatively small, reduced *EFEMP1* expression appears to be a frequent and is also a subtype-independent. However, an interesting observation is that where increased expression is apparent, it appears to be confined to the corticotrophinoma. The number of subtype-specific adenomas, showing reduced expression of *EFEMP1*, are summarised in Table 5.1.

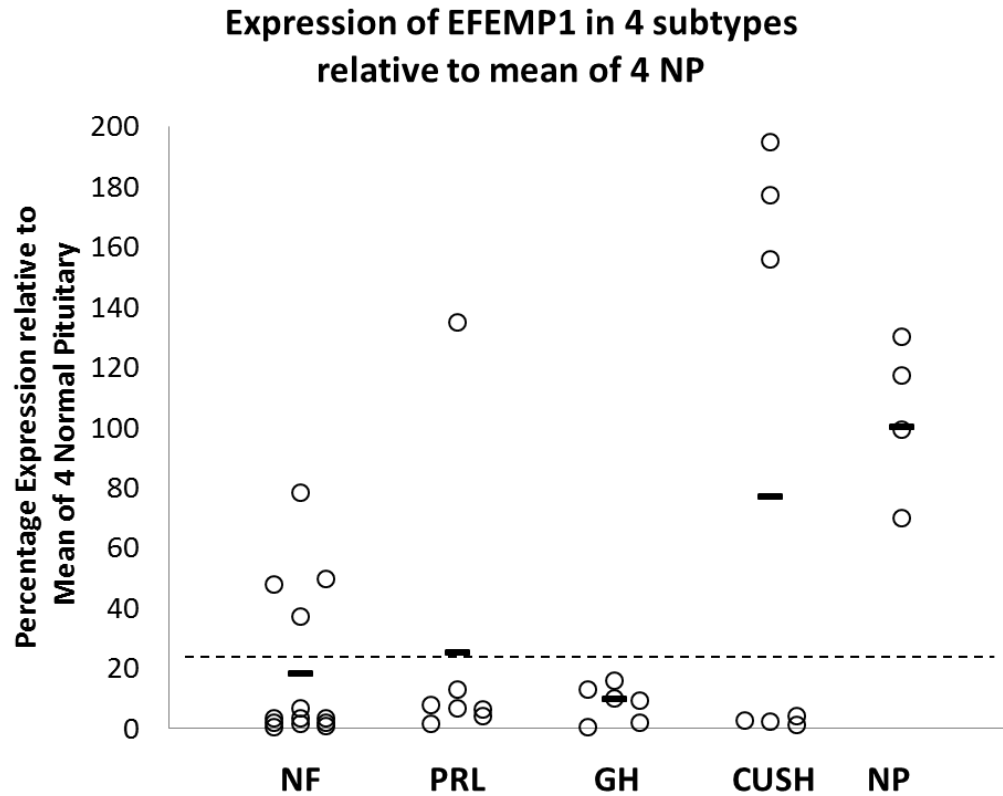


Figure 5.1. Quantitative RT-PCR analysis of *EFEMP1* expression in each of the major pituitary adenoma subtypes.

Expression is shown relative to four post-mortem normal pituitaries (NP). Adenoma were assigned as showing significantly reduced expression where expression was ≥ 3 SD, lower than the mean of expression in the NP (dotted line). The horizontal bars show the mean expression in each adenoma subtype and in NP. Small unfilled circles are individual adenoma. Non-functioning adenoma (NF), prolactinoma (PRL), growth-hormone secreting adenoma (GH) and corticotrophinoma (CUSH). Relative to the housekeeping gene *GAPDH*, mean expression level of *EFEMP1* in 4 NPs, as estimated by absolute Ct values in RT-qPCR, is approximately 2%.

Subtype	Number of sample	Reduced expression	Normal expression
NP	4	0	4
NF	13	9	4
PRL	7	6	1
GH	6	6	0
CUSH	7	4	3

Table 5.1. *EFEMP1* expression status in pituitary adenoma.

The number of adenomas that express or show reduced expression of *EFEMP1*. Expression status as determined and relative to post-mortem normal pituitaries using the criteria described in the materials and methods.

5.2. Methylation status of the *EFEMP1* genes CpG island in pituitary adenomas

I next determined the methylation status of the *EFEMP1* gene's CpG island in the investigation cohort of adenoma and relative to post-mortem pituitaries. Pyrosequencing analysis showed that methylation (on the basis of the criteria described in the materials and methods section), was not confined to GH secreting adenomas but is also apparent in each of the other adenoma subtypes. In this case, the segregation with the GH adenoma was on the basis of the original array findings of the investigation cohort and shown in Figure 4.2 (heat map). An example of methylation as determined by pyrosequencing is shown in Figure 2.1 and the findings are summarised in Table 5.2. In these cases, mean methylation across four normal pituitaries was 15.5% and adenomas were classed as methylated in those cases where the mean was ≥ 3 or more times the S.D. of that mean, that is, $\geq 41.7\%$. The apparent exception are prolactinoma, however, in this and several other instances, not all of the adenomas available for expression analysis yielded useable products for pyrosequence analysis following sodium bisulphite conversion. Table 5.2 shows, in summary, that none of the adenoma that expressed *EFEMP1* are methylated. However, in

tumours that do not express *EFEMP1*, approximately equal proportion show this epigenetic change as those that do not. I concluded that, although methylation is not seen in *EFEMP1* expressing adenoma or normal pituitaries neither the absence or the presence of this change “predicts” adenomas that failed to express this transcript. These findings suggested that mechanisms in addition to CpG island methylation are responsible for silencing of *EFEMP1*.

Subtypes	Number of samples	<i>EFEMP1</i> Expressing Adenomas		<i>EFEMP1</i> Non-Expressing Adenomas	
		Methylated	Unmethylated	Methylated	Unmethylated
NP	4	0	4	0	0
NF	13	0	4	5	4
GH	5	0	0	2	3
CUSH	3	0	2	1	0
PRL	3	0	1	0	2

Table 5.2. *EFEMP1* expression status in pituitary adenoma relative to CpG island methylation status.

The number of adenomas that either express or do not express *EFEMP1* are shown relative to their methylation status. The criteria for methylation and expression are described in the materials and methods section.

5.3. Histone modifications associated with *EFEMP1* expression in primary adenoma

The influence of histone modification on *EFEMP1* expression was assessed by ChIP analysis. Primer design was described in section 2.7.5. In these cases it was necessary to confine these studies to the non-functioning adenomas due to absolute tumour size and the reliability of chromatin extractions. However, clear trends in enrichment and depletion are evident. Figure 5.2 shows that in adenomas showing reduced *EFEMP1* expression the histone modification, H3K9Ac is depleted and the H3K27me3 is enriched relative to post-mortem normal pituitaries. In adenomas that express *EFEMP1* the converse was apparent, where adenomas were enriched for the H3K9Ac modification and depleted for H3K27me3. Furthermore, for two of adenomas (NF7 and NF10), reduced expression of *EFEMP1* is associated with histone modifications, however, these adenomas did not show change in CpG island methylation. These findings show that histone modification associate with active and repressed genes are associated with the expression status of *EFEMP1* in primary pituitary adenoma. Moreover, the histone modifications associated with reduced *EFEMP1* expression is not invariably accompanied by promoter-associated CpG island methylation.

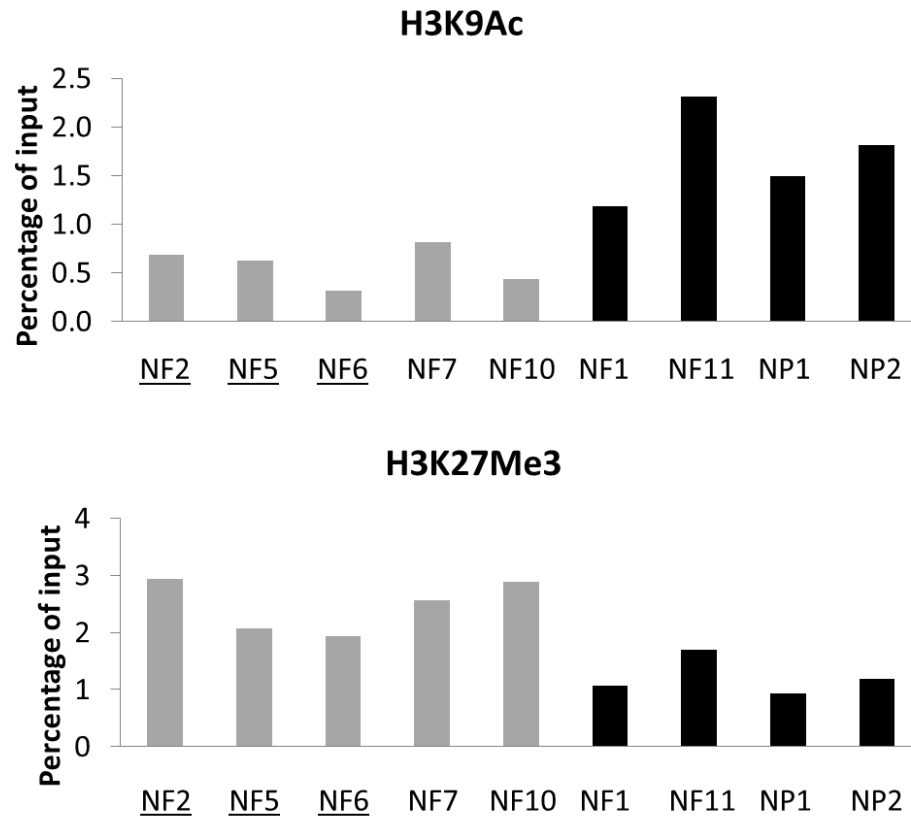


Figure 5.2. ChIP Analysis of the *EFEMP1* gene.

The upper panel shows ChIP enrichment for the H3K9Ac modification and lower panel for the H3K27me3 modification. Enrichments are shown for non-functional adenoma (NF) and post-mortem normal pituitaries (NP). All experiments, in available adenomas, were performed at least three times and representative examples shown for each. Due to limitation of available tumour specimen it was not possible to examine all of the adenomas shown in Figure 5.1. Due to small size of tumours it was not possible to perform three replicates, therefore only one ChIP was performed per tumour. An underline, of an individual adenoma, signifies an adenoma that did not show methylation of *EFEMP1* within its CpG island. Primers targeted onto the position of -33 to +87 relative to TSS.

5.4. Epidrug induced re-expression of the *Efemp1* gene in AtT20 and GH3 cells

The murine and rodent pituitary cell lines, AtT20 and GH3 respectively express *Efemp1* at marginal levels as determined by RT-qPCR. In contrast to these findings their cognate normal pituitary counterparts show robust expression of the *Efemp1* transcript (Figure 5.4). In these cell lines, combined drug challenges, with zebularine and TSA, induced re-expression of *Efemp1*. In cases of pituitary cell lines GH3 and AtT20, zebularine instead of 5-Aza-2-Dc was chosen as an epi-drug because of our lab's previous experience regarding 5-Aza-2-Dc cytotoxicity to these cell lines. Zebularine was reported by others to be a less toxic demethylating agent (Cheng *et al.*, 2003; Cheng *et al.*, 2004b). Single drug challenge showed marginal or no effect on transcript expression, however, increasing doses of zebularine, and where TSA concentrations were constant, show a dose dependent increase in *Efemp1* as determined by RT-qPCR.

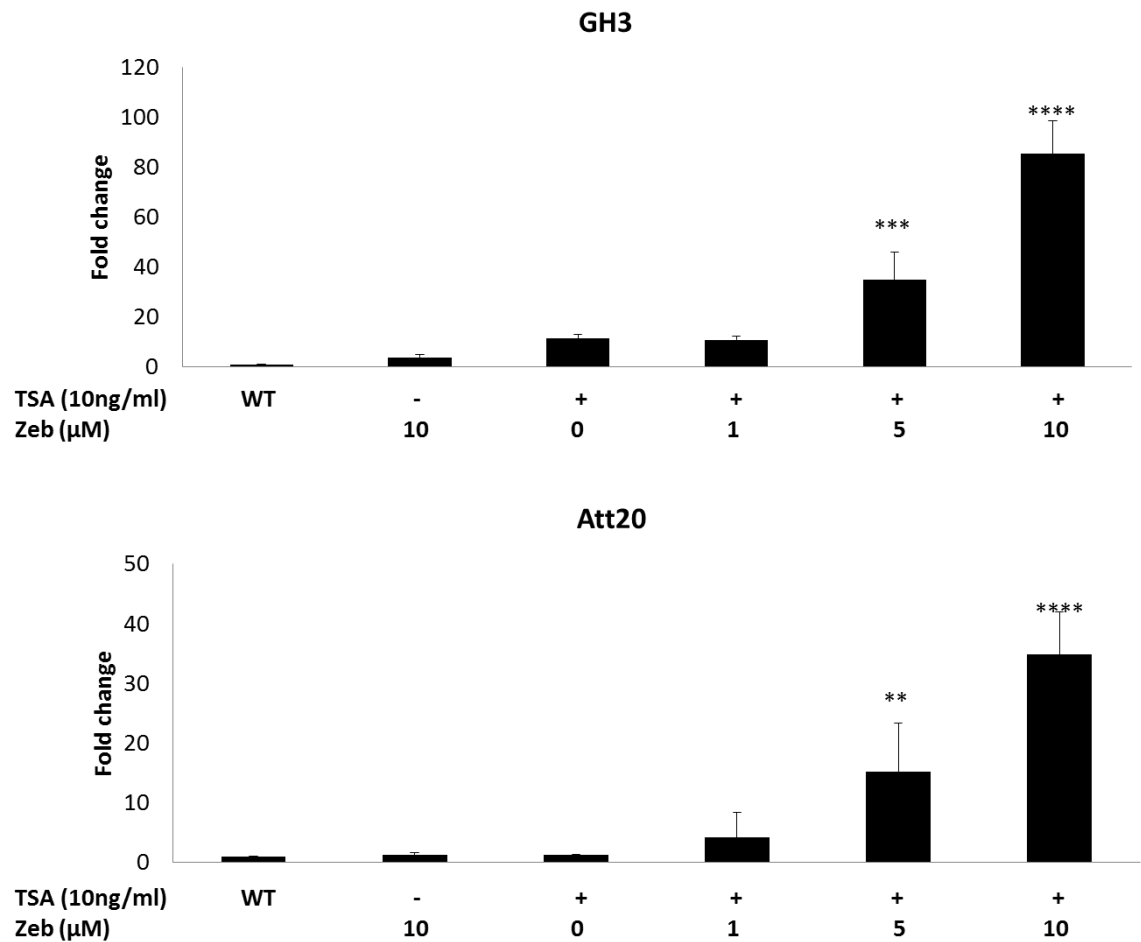


Figure 5.3. Expression status of *Efemp1* in pituitary cell lines.

Upper panel, GH3 cells and lower panel, AtT20 cells following challenge with zebularine (ZEB) or TSA alone or in combined incubations. In each case, the panels show RT-qPCR analysis of *Efemp1* in response to the challenges and drug doses shown in the figures. All experiments were done three times with triplicate determinations in each experiment. Error bars show SEM. **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$ vs. vehicle alone. Data were analysed for significance by one-way ANOVA with Dunnett's multiple comparison posttest. Relative to the housekeeping gene *Pbgd*, expression levels of *Efemp1* in NRP and NMP, as estimated by absolute Ct values in RT-qPCR, are approximately equal and slightly higher, respectively. NRP, Normal rat pituitary; NMP, normal mouse pituitary; WT, wild type.

5.5. Epigenomic changes associated with expression of *Efemp1* in pituitary cell lines

In both the AtT20 and GH3 cell lines, silencing of the *Efemp1* gene is associated with a significant increase in CpG methylation as determined by pyrosequencing and within their respective CpG islands. In these cases increase was determined relative to their cognate normal pituitaries (Figure 5.4). On the basis of these findings I reasoned that gene silencing and epidrug induced re-expression was consequent to CpG island methylation. However, analysis of cell lines following the drug treatment described in Figure 5.3 did not show change in DNA methylation in either AtT20 or GH3 cells.

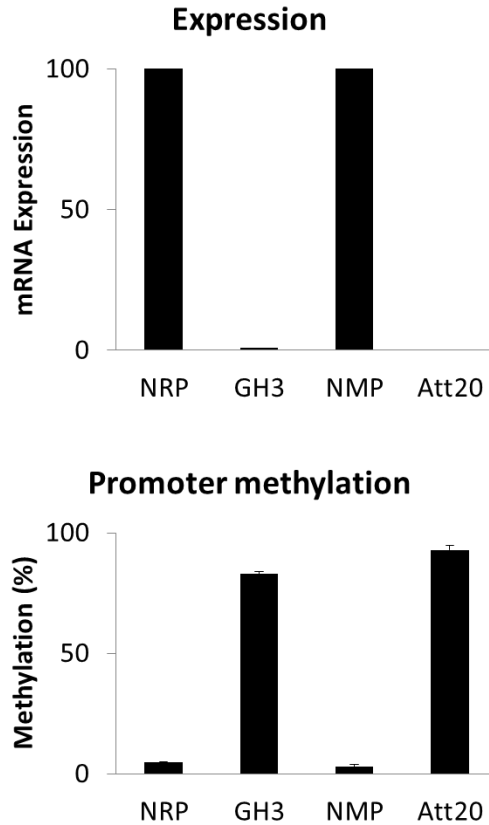


Figure 5.4. Inverse correlation between transcriptional expression and promoter methylation of *Efemp1* in rat and mouse pituitary models.

Expression at transcriptional level of *Efemp1* (upper panel) in both GH3 and Att20 cell lines is significantly reduced relative to normal rat and normal mouse pituitary, respectively. Low expression levels in cell lines are correlated with high methylation levels of promoter-associated CGI (lower panel) while the reverse phenomenon is apparent in normal rat and mouse pituitary. Relative to the housekeeping gene *Pbgd*, expression levels of *Efemp1* in NRP and NMP, as estimated by absolute Ct values in RT-qPCR, are approximately equal and slightly higher, respectively. NRP, Normal rat pituitary; NMP, normal mouse pituitary. 142.

5.6. Histone modifications in epidrug challenged pituitary cell lines

Previous reports from our laboratory have shown that epidrugs induce histone modifications and that these changes are associated with re-expression of silenced genes (Al-Azzawi *et al.*, 2011; Yacub-Usman *et al.*, 2012a). We therefore investigated histone modifications associated with the promoter region of the *Efemp1* gene in AtT20 and GH3 cells. Figure 5.5 shows that combined epidrug challenges leads to increase, which is clearly in GH3 and slightly in AtT20, in the enriched fraction of the H3K9Ac modification. Conversely, in both cell lines drug challenges led to a decrease in the modification commonly associated gene silencing, H3K27me3. These findings suggest that histone modification and specifically enrichment of the H3K27me3 mark and depletion of the H3K9Ac modification are responsible for silencing of the *Efemp1* gene in AtT20 and GH3 cells.

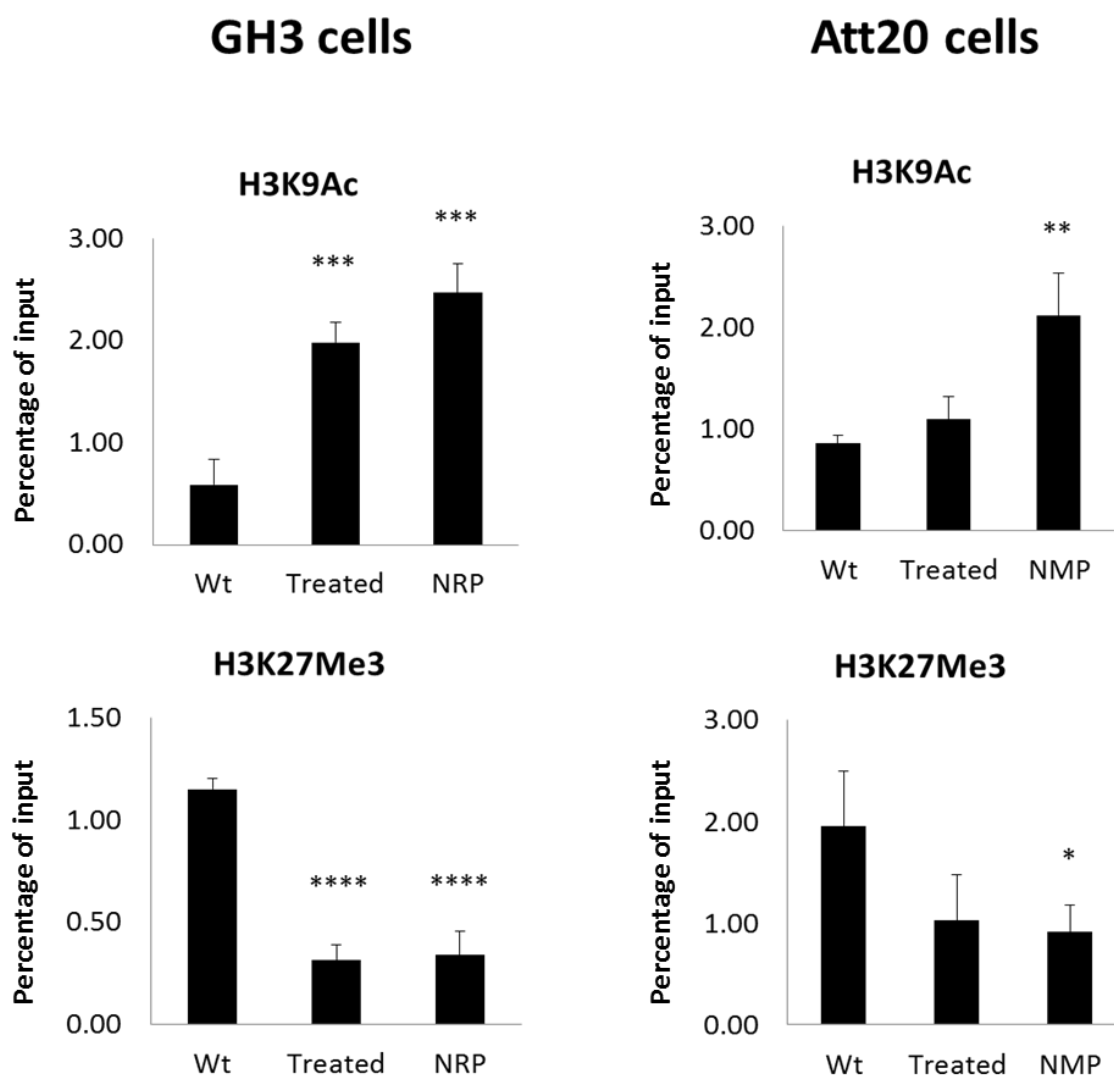


Figure 5.5. ChIP analysis of *Efemp1* promoter region in pituitary cell lines.

Analyses are shown relative to the modification seen in input DNA for the H3K9Ac and H3K27Me3 Enrichment in the pituitary cell line GH3 and AtT20, left and right side of the figure respectively. Enrichment is shown following drug challenge (treated) or vehicle control (WT) shown in figure 3 and relative to the enrichment observed in either, normal rat (NRP) or normal mouse (NMP) pituitary. All experiments were done three times with triplicate determinations in each experiment. Primers for rat and mouse targeted onto the position of +77 to +175 and +159 to +221 relative to TSS, respectively. Error bars show SEM. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$ vs. vehicle alone. Data were analyzed for significance by one-way ANOVA with Dunnett's multiple comparison posttest. NRP, Normal rat pituitary; NMP, normal mouse pituitary; WT, wild type.

5.7. Function end-points contingent on heterologous *EFEMP1* expression in GH3 cells

To investigate putative functional end points, stable transfected GH3 cells harbouring and inducible human *EFEMP1* construct was first generated. Figure 5.6 shows *EFEMP1* expression as determined by RT-qPCR in the presence and absence of the inducing agent, IPTG. In these studies, a dose dependent increase in *EFEMP1* expression between 5-100 μ M was observed.

5.7.1. Proliferation and apoptotic end-points

The effects of enforced *EFEMP1* expression (100 μ M IPTG) on cell proliferation and apoptosis were assessed in GH3. Relative to GH3 harbouring an empty vector control or the *EFEMP1* construct challenged with vehicle alone, no change in proliferation was apparent in a 12 day growth curve. In studies designed to assess inhibition or augmentation of apoptotic response we used bromocriptine challenge and caspase 3/7 activation and as we have described previously (Rowther *et al.*, 2010). In these experiments and similar to those described for proliferative responses I did not detect an effect on this end point (data not shown).

5.7.2. *EFEMP1* target gene end-points

The effects of enforced *EFEMP1* expression in GH3 cells and as described above were determined for genes shown by others to be subject to the effects of enforced *EFEMP1* expression. In these cases, *EFEMP1* expression had no discernible effects on the expression of the endogenous genes coding for *VEGF-A*, *HIF1A*, *MMP7* or *MMP9* (data

not shown). However, a dose dependent inverse relationship between enforced *EFEMP1* expression and that of *MMP2* was observed (Figure 5.6).

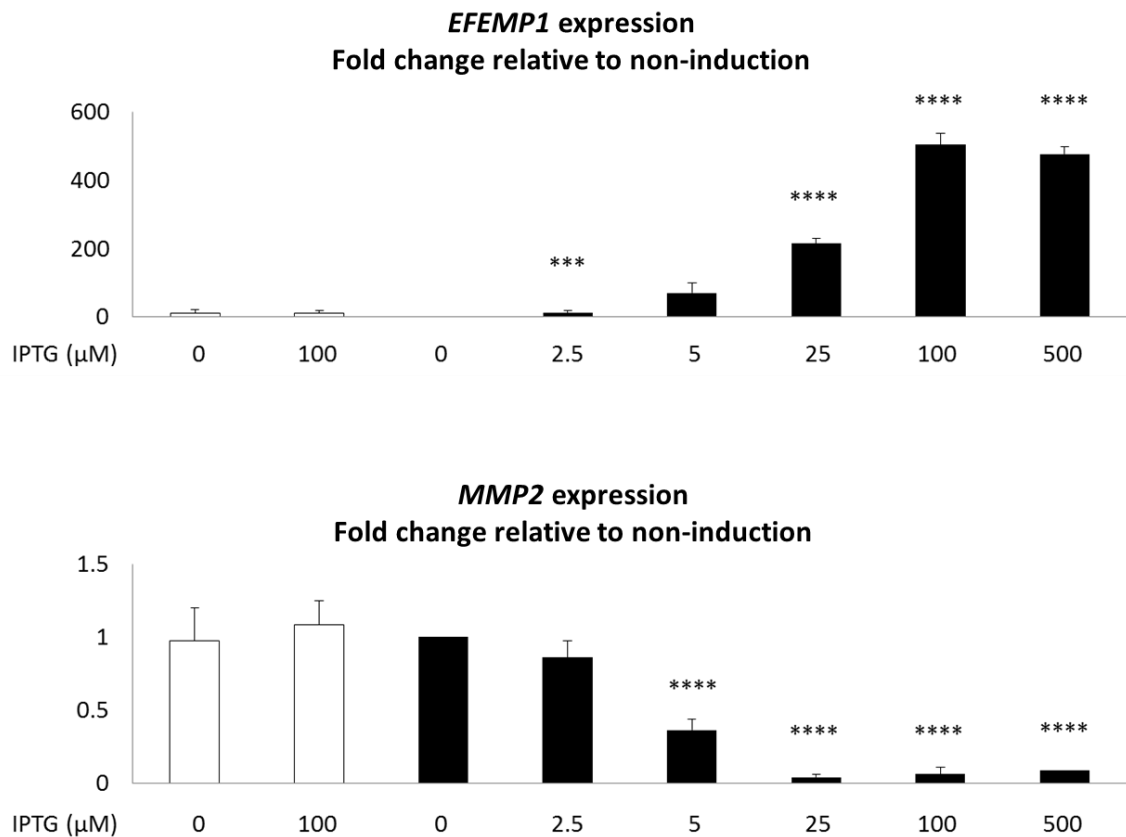


Figure 5.6. *MMP2* expression in stable transfected GH3 cells.

Upper panel shows *EFEMP1* expression induced by varying doses of IPTG and as determined by RT-qPCR. Lower panel shows *MMP2* expression in corresponding cells indicated in the upper panel. In each case the fold induced is shown relative to cells challenged with vehicle alone. All experiments were done three times with triplicate determinations in each experiment. Error bars show SEM. ***, $P < 0.001$; ****, $P < 0.0001$ vs. vehicle alone. Data were analyzed for significance by one-way ANOVA with Dunnett's multiple comparison posttest. Open bars are control experiments performed on cells transfected by empty vector.

Finally, in the pituitary adenoma cohort used in this study I determined the relationship, if any, between *EFEMP1* expression and the target genes described above. Using the cut-off criteria described above we found a positive relationship between *EFEMP1* expression and that of *MMP7* (Figure 5.7), however, for the others genes I investigated in GH3 cell I did not see this relationship.

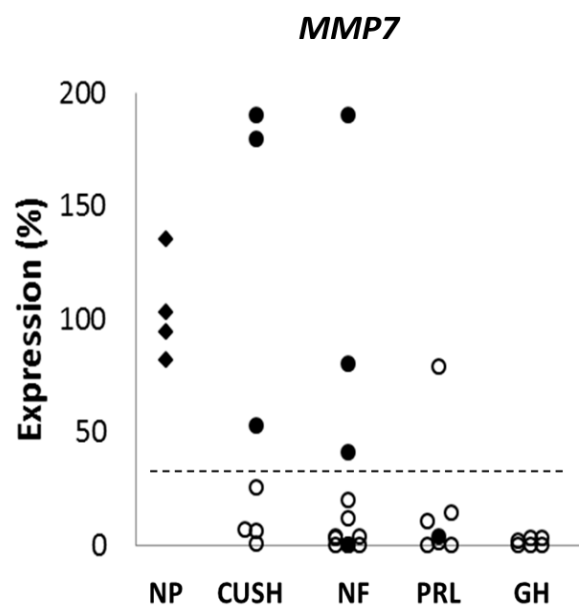


Figure 5.7. Expression of *MMP7* in pituitary adenoma in the absence and presence of *EFEMP1*.

The dotted line is three SD less than the mean expression of *MMP7* in post mortem pituitaries (diamonds). For each of the adenoma subtypes, *MMP7* expression is plotted relative to the average *MMP7* expression in four NP. In these cases, filled circles are adenoma that express *EFEMP1* and unfilled circles are adenoma that show loss or significantly reduced *EFEMP1* expression. Abbreviations are as shown in the legend to Figure 5.1.

5.8. Discussion

Recent years have witnessed considerable and rapid advances in the technologies applicable to the analyses of the genome and of the epigenome in normal and disease states (Emes and Farrell, 2012). Furthermore, our understanding of changes to the epigenome and that impact on gene expression has burgeoned (Emes and Farrell, 2012; Yacqub-Usman *et al.*, 2012b). We now know that DNA methylation, histone modifications and miRNA expression patterns, either alone or more usually in concert impact on gene expression and translation (Dudley *et al.*, 2009; Yacqub-Usman *et al.*, 2012b). In addition to the repertoire of modification, their consequences at the transcript or protein level are frequently apparent in a cell- or tissue- specific context (Hayward *et al.*, 2001; Farrell and Clayton, 2000; Simpson *et al.*, 2000; Fedele *et al.*, 2006).

In this thesis, I focused attention on the *EFEMP1* gene where Illumina BeadArray analysis had demonstrated differential CpG island methylation confined to GH secreting adenomas relative to the other major adenoma subtypes and to that apparent in post-mortem normal pituitaries (Table 4.2). However, in an independent cohort of pituitary adenomas I now show by pyrosequencing analysis that methylation of the *EFEMP1* gene's CpG island is not confined to GH secreting adenomas but is also detected in the other adenoma subtypes. Our failure to detect this change in a previous investigation cohort is most likely a consequence of only three, of each adenoma subtype, being interrogated on the BeadArray. Moreover, I now show that this epigenetic modification, if considered in isolation, is not invariantly associated with loss or reduced *EFEMP1* expression. Indeed, although the majority of adenomas showed barely detectable *EFEMP1* transcript expression approximately equal numbers of adenoma showed CpG methylation as did not show this change. Our research group and others have made this observation in previous

studies where CpG island methylation is variably associated with gene silencing (Duong *et al.*, 2012; Kim *et al.*, 2011a; Zeller *et al.*, 2012).

These observations, of gene silencing in the presence and absence of CpG island methylation suggested alternate mechanisms of gene silencing. I, therefore, determine the influence of histone modifications, as a potential independent or co-dependent contributor toward gene silencing. For those tumours where sufficient sample was available ChIP analysis showed enrichment for the histone modification, H3K27me3 and depletion of the H3K9Ac modification associated with reduced expression of *EFEMP1*. Enrichment for these histone tail modifications, as shown by several groups, including our own, are associated with active (H3K9Ac) and repressed (H3K27Me3) gene expression (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a; Cao *et al.*, 2002b; Jenuwein and Allis, 2001). These conclusions were reinforced by the contrary findings, where enrichment (H3K9Ac) and depletion (H3K27me3), in tumours and post-mortem pituitaries that expressed the *EFEMP1* gene, is observed. Indeed, in a previous studies our research group observed that significantly reduced expression of BMP-4, in pituitary adenoma, is also associated with the histone modifications that I now described for the *EFEMP1* gene (Yacqub-Usman *et al.*, 2012a).

Since multiple studies show that epimutations, apparent as CpG methylation and/or as histone modifications are potentially reversible (Yacqub-Usman *et al.*, 2012b; Sharma *et al.*, 2010) I adopted these types of approach in the pituitary tumour cell lines, GH3 and AtT20. Both cell lines expressed *Efemp1* at marginal levels but in contrast to human pituitary tumours, their gene associated CpG islands were heavily methylated. However, in common with human pituitary tumours they also showed histone modification indicative of gene silencing. Combined incubations with the epidrugs, zebularine and TSA led to a dose dependent increase in *Efemp1* as determined by RT-qPCR. In these cases, epidrugs-

mediated expression did not induce or lead to change in the methylation status of the CpG islands associated with the *Efemp1* gene. However, the epidrug induced histone modifications that are indicative of and associated with active gene and these modifications were more marked in GH3 cells than apparent in AtT20 cells. In previous studies of epidrug challenged cells, in these cases for the D2R and BMP-4 gene we saw similar, marked changes in histone modifications (Al-Azzawi *et al.*, 2011; Yacqub-Usman *et al.*, 2012a). Moreover, in these cases these changes were also accompanied by a decrease, albeit marginal, in CpG island methylation. While the differences in drug-induced changes to the epigenome are not entirely clear we did note that, although the increase in *Efemp1* was significant the absolute levels, as determined by RT-qPCR, were significantly lower than that seen in their normal pituitary counterparts (data not shown).

The *EFEMP1* gene encodes fibulin-3 and is one of seven members of the fibulin family of extracellular matrix proteins (ECM) (Lecka-Czernik *et al.*, 1995). These multifunctional proteins are located in basement membranes, stroma, and ECM fibres and their functions include cell-to-cell and cell-to-matrix communication, organogenesis and vasculogenesis and the organization and stabilization of ECM structures (de Vega *et al.*, 2009). In the context of tumorigenesis fibulins impact on multiple pathways that include tumour suppressor, apoptotic, angiogenic, metastatic, and invasion (Gallagher *et al.*, 2005; Obaya *et al.*, 2012; Seeliger *et al.*, 2009). Moreover, the contribution of some family members, as either promoters or inhibitors of these pathways is frequently apparent in a cell- or tumour- type context (Obaya *et al.*, 2012). As example, several studies show that *EFEMP1* has *bona fide* activity as either a tumour promoter or tumour suppressor (Seeliger *et al.*, 2009; Albig *et al.*, 2006; Hu *et al.*, 2011; Song *et al.*, 2011b). Indeed, increased expression of *EFEMP1* is apparent in pancreatic and cervical cancers and malignant glioma, (Song *et al.*, 2011b; Hu *et al.*, 2009; Seeliger *et al.*, 2009), whereas, reduced

expression is apparent in breast, prostate, hepatocellular, and non-small-cell lung cancer (NSCLC) (Kim *et al.*, 2011b; Sadr-Nabavi *et al.*, 2009; Nomoto *et al.*, 2010; Kim *et al.*, 2012). The mechanisms responsible for inappropriate expression of *EFEMP1*, and in particular for increase are poorly understood. However in multiple tumour types, including NSCLC, breast, prostate, and hepatocellular cancer loss or reduced expression is associated with CpG island methylation (Sadr-Nabavi *et al.*, 2009; Kim *et al.*, 2011b; Nomoto *et al.*, 2010; Kim *et al.*, 2012). Indeed, and similar to our findings in GH secreting adenoma Kim and colleagues employed a BeadArray approach and identified increase methylation of the *EFEMP1* gene in prostate cancer (Kim *et al.*, 2011b).

The functional consequences of enforced *EFEMP1* expression were determined in GH3 cells harbouring an inducible construct. The end-points measured included effects of *EFEMP1* expression on cell proliferation, apoptosis and expression of *EFEMP1* responsive genes. In these cases and for these end-points several studies have shown *EFEMP1* mediated effects in different cell and tumour types (Kim *et al.*, 2012; Song *et al.*, 2011b; Seeliger *et al.*, 2009; Hu *et al.*, 2011; Albig *et al.*, 2006). Enforced expression of *EFEMP1* in GH3 cells did not inhibit their proliferation or drug-induced apoptotic response, whereas, in NSCLC, pancreatic, and glioma tumours *EFEMP1* inhibits one or more of these functional endpoints (Kim *et al.*, 2012; Seeliger *et al.*, 2009; Hu *et al.*, 2012). Our findings in pituitary cells might, therefore, reflect the fact that these functional end points represent cell- or tumour- type specific phenomenon, as observed by other investigators (Song *et al.*, 2011b; Seeliger *et al.*, 2009; Hu *et al.*, 2011). Indeed and in this context, enforced expression of *EFEMP1* in Hela cell, promotes, and in NSCLC cells inhibits proliferation (Kim *et al.*, 2012; Song *et al.*, 2011b) and similar finding of an association between growth promotion or growth inhibition in xenograft models of

pancreatic tumours and glioma respectively are also reported (Seeliger *et al.*, 2009; Hu *et al.*, 2011).

The functional consequences of *EFEMP1* expression on tumour invasion and angiogenesis also show tumour type specificities, where expression may either promotes or inhibits these endpoints (Albig *et al.*, 2006; Hu *et al.*, 2009; Hu *et al.*, 2011; Song *et al.*, 2011b). The impact of *EFEMP1* on these end-points is frequently mediated through effects on expression of multiple gene targets, including *VEGF-A*, *HIF1A* and *MMPs*. As example, overexpression of *EFEMP1* leads to increase in *VEGF-A* expression in pancreatic and cervical cancer cells (Seeliger *et al.*, 2009; Song *et al.*, 2011b) and to decrease of this growth factor in malignant glioma cells (Hu *et al.*, 2011). For the *MMPs*, a positive correlation is apparent between *EFEMP1* and *MMP2* and *MMP9* expression in glioma, and enforced expression of *EFEMP1* in a cell line model also up-regulates these genes (Hu *et al.*, 2009). Conversely, in NSCLC cells line, enforced expression and siRNA knock-down of *EFEMP1* leads to decreases or increase of *MMP2/7* respectively (Kim *et al.*, 2012). However, although our studies of enforced *EFEMP1* expression in GH3 did not show change in the expression of either, *VEGF-A* or that of *HIF1A* I did observe a dose dependent, *EFEMP1* mediated, decrease in *MMP2* expression in these cells. Finally, I looked for correlation between *EFEMP1* expression and that of *VEGF-A*, *HIF1A* and *MMPs* in the cohort of primary human tumours. A positive correlation between *EFEMP1* and *MMP7* was apparent, however, no other correlation or associations were apparent across the tumour cohort. The lack of an association between *EFEMP1* and *MMP2*, apparent in GH3 cells but not in human tumours, might reflect species or tumour type difference in our studies.

In conclusion, my thesis shows that the majority of pituitary adenomas show significantly reduced expression of the multi-functional protein fibulin-3 encoded by the

EFEMP1 gene. I show for the first time that loss is subtype independent and is frequently but not invariably associated with CpG island methylation and or histone modification. For histone modifications this change may lead to silencing that is independent of CpG island methylation. Epidrug challenges lead to re-expression of *EFEMP1* in a cell line model and show a functional relationship between histone modification and gene silencing. Of several endpoints investigated, through enforced expression of *EFEMP1*, an inverse relationship was apparent between *EFEMP1* and *MMP2*; however, this relationship was not apparent in primary human tumours and might represent species or cell type differences. The positive relationship between *EFEMP1* and *MMP7* expression in human tumours is intriguing and warrants further investigation.

Chapter 6: Conclusions and future works

6.1. Conclusions

The introduction to this thesis has provided an overview of pituitary tumourigenesis and described the genetic and epigenetic changes/aberrations associated with this tumour type. Furthermore, with respect to the balance between genetic and epigenetic changes the literature suggests that the latter are more frequent than the former (Yacqub-Usman *et al.*, 2012b). Moreover, most of these changes, that are epigenetic, have been identified by candidate gene approaches. In those infrequent cases where novel genes have been identified the identification has been constrained, by technical limitations to identification on a gene-by-gene approach. These inherent limitations have also been compounded by the small size of these adenomas and, therefore, limited material, be it DNA, RNA or protein for down-stream analysis. In this final chapter I will review the progress made toward surmounting these limitations and also described potential future studies that address some of the questions raised by these new findings.

The initial aim of this thesis was to profile DNA methylation at genome-wide level within each of the four major pituitary adenoma subtypes. For this purpose, I first need to overcome the technical limitation imposed by limited primary tumour material. Therefore, in Chapter 3, I described the optimisation and validation of WGA as applied to sodium bisulphite converted DNA. I was able to show that WGA is reliable and reproducible in the context of quantitative amplification and also with respect to qualitative aspects. Quantitatively, it provides a reliable and consistent 25 fold amplifications of sodium bisulphite converted DNA. Qualitatively, there is no discernible bias in the DNA methylation profile between DNA prior to WGA and that after derived after this

amplification process. Interestingly, although not part of published data I found that whole genome amplified DNA could be stored for longer periods than converted DNA and that primers directed to converted DNA provided more consistent amplification than those used on DNA prior to amplification.

The materials derived from WGA were interrogated using the Illumina Infinium 27K assay that interrogates 27,578 CpG sites across 14,495 genes. In Chapter 4, I first described how this genome-wide study was designed. In this case, the 27K assay was performed on 15 pituitary tumours (discovery cohort) that included 3 samples of each pituitary adenoma subtype (NF, GH, PRL, and CUSH) and relative to three post-mortem pituitary glands. The data set generated by 27K BeadArray assay was processed *in silico* using various exclusion criteria, initially for quality control and subsequently by exclusion of all CpGs that were regarded to be lower than or at the limits of detection (≤ 0.2 [β value]) or showed values ≥ 0.8 across the data set. The resulting data set (after exclusion of data from the X-chromosome) resulted in 10,667 CpGs spanning 7,956 genes. Second, the reliability of the BeadArray-derived data was assessed through validation experiments using sodium bisulphite pyrosequencing. The results section describes the validation and shows that the majority of array derived β values were confirmed using an independent technique, in this case pyrosequencing. Having established the validity and feasibility of the genome-wide approach it was possible to identify genes that are hypermethylated in pituitary tumours relative to post mortem normal pituitaries. In this case preliminary studies showed that it was necessary to use a stringent cut off criteria where one of two CpGs showed a $\Delta\beta$ of ≥ 0.4 and a second CpG a $\Delta\beta$ of ≥ 0.25 . This identified 40 genes in NF adenomas, 21 in GH, 6 in PRL and 2 in CUSH. Of these, I randomly selected 22 genes for promoter methylation analysis by pyrosequencing using, in this case a larger investigation cohort of pituitary tumours. Of the 22 genes investigated I was able to

confirm methylation in 16 cases, that is 75% and relative to normal pituitaries. Of the 16 genes studied in the investigation cohort different frequencies of methylation were observed. Moreover, and interestingly, 3 genes showed subtype-specific DNA methylation patterns. To determine potential associations between methylation and gene expression I used RT-qPCR. In these cases a causal relationship between methylation and expression was only observed for three of these genes. In these cases a likely explanation is the “driver-passenger” theory (Zeller *et al.*, 2012; Kim *et al.*, 2011a) and/or multiple epigenetic mechanisms where mechanisms other than or in addition to promoter hypermethylation are responsible for gene inactivation. In the final results chapter of my thesis I addressed this phenomenon by looking at histone modification for one of the identified genes (see below). Finally, I performed gene ontology analysis and found that a number of hypermethylated genes identified in my BeadArray analysis had also reported as tumour suppressor genes in other types of tumours. The output of my genome-wide study provides the first, subtype-specific, identification of candidate genes probably involved in pituitary tumorigenesis for further studies. Moreover, while candidate gene studies frequently show an association between CpG island methylation and loss or significantly reduced gene expression the BeadArray analysis did not. We presume this reflects the following: assuming reduced transcript levels are first shown then the likelihood of finding this to be associated with methylation is significant. However, where the initial part of the investigation is methylation status then the presence of this change is not invariably associated with loss of transcript. Our conclusion would be that the presence of increased methylation “suggests” that a gene might be transcriptionally incompetent but does not establish a significant association.

To extend my studies toward a functional analysis of identified hypermethylated genes, I focused attention on the *EFEMP1* gene. In part the rationale for investigating this

gene was that it appeared to show subtype-specific hypermethylation that was confined to the GH secreting adenomas. Furthermore it has been reported as multi-functional protein in different tumour types and show context specific actions. Somewhat surprisingly I found that expression of *EFEMP1* at transcriptional level is significantly reduced in 25 of 33 pituitary tumours (~75%) and loss or reduced expression was irrespective of pituitary adenoma subtype. To investigate mechanisms responsible for reduced expression, I first examine DNA methylation of the promoter-associated CpG island of the *EFEMP1* gene by pyrosequencing. Due to sample limitations, only 28 of 33 samples were available for this analysis. While these studies showed that adenomas that expressed *EFEMP1* were not methylated (relative to normal pituitary) a more complex picture was apparent for tumours showing reduced expression. In these cases, of the 17 adenomas that showed reduced expression of *EFEMP1* only 8 were methylated. Therefore, in some tumours methylation could not account for reduced expression.

To gain further insight into mechanisms associated with expression of the *EFEMP1* gene, I extended studies to rat and mouse pituitary cell line models. In both species, significantly reduced expression of *Efemp1* was associated with CpG island methylation and histone modifications. In human pituitary adenomas I also showed a significant association between histone modification and gene silencing. Taken together, ChIP data clearly showed a correlation between histone modifications and *EFEMP1* gene expression. However, additional experiments were needed to examine the causal effects of DNA methylation and/or histone modifications on *EFEMP1* gene expression. Toward this goal, I treated rat and mouse pituitary cell lines with epigenetic drugs either alone or in combination. In these cases, co-incubation with the epigenetic drugs restored expression of this transcript. Restored expression was associated with increased and decreased enrichments of histone modifications indicative of active- and inactive-gene expression, respectively.

Taken together, my data determined a causal effect of histone modifications on *EFEMP1* gene expression.

Finally, for functional analysis I used enforced-expression of *EFEMP1* in rat pituitary cells by using an inducible expression vector. The results showed that this gene had no effects on cell proliferation or apoptotic end points. However, a dose-dependent inverse effect *EFEMP1* and *MMP2* was apparent. Interestingly, in human pituitary samples this observation was not seen, but there was a direct correlation between *EFEMP1* and *MMP7*. These initial observations require further studies. However, the data gained within this thesis added to the literature in which fibulin-3 showed different behaviours dependent on cell types.

6.2. Suggestions for future work

6.2.1. Up-to-date genome-wide DNA methylation profiling of pituitary tumours

Genome-wide DNA profiling technologies have made significant advances. At the start of my PhD studies the Illumina 27K BeadArray was the most technically advanced means of investigating the genome on an genome-wide basis. Subsequent to these studies arrays that interrogated more than 450,000 CpG sites have been described and are commercially available. In addition there are a number of commercially available Next Generation Sequencing technologies and other array-based genome-wide DNA methylation strategies available. Some of these techniques and technologies are reviewed in section 1.3.2.II of this thesis and may form avenues for further studies in this tumour type. A limitation of my studies was also apparent and due to the limited number of samples interrogated for each adenoma subtype. Therefore, I would like to perform another

genome-wide study in which a larger number of pituitary samples is interrogated at higher coverage ratio of epigenome.

In this thesis, I focused on CpGs located within CpG islands where I saw increased methylation relative to normal pituitaries. In this way an aim was to identify putative tumour suppressor genes. However, there are some studies of oncogenes reported to be over-expressed in tumours relative to normal tissues through promoter hypomethylation (De Smet *et al.*, 1996; Singh *et al.*, 2003). In these cases it is necessary to identify hypomethylated CpGs. Using the similar stringent cut-off criteria applied to identify hypermethylated CpGs, screening of hypomethylated CpGs did not return any result. This was likely because of the limitation of the number of samples interrogated. Therefore, in the next 450K genome-wide investigation I suggest it will be important to give equal attention to the finding of hypomethylation genes. Moreover, recent reports showed that methylation of CpG sites located outside of CpG islands is also important (Doi *et al.*, 2009; Irizarry *et al.*, 2009; Portela and Esteller, 2010). Thus, I am keen to investigate aberrantly methylated CpG sites located within CpG island shores.

6.2.2. Further investigation of the candidate genes identified in this thesis

One of the major outputs of my thesis is the list of genes which have their promoter-associated CpG islands hypermethylated in tumours relative to normal pituitaries. One of them, *EFEMP1*, had been investigated in the later part of this thesis. The results indicated that the rest of this list can potentially be a valuable resource of candidate genes for further studies in pituitary tumourigenesis research. Attention probably should focus on 16 genes whose aberrant methylation status between tumours and normal pituitaries had been confirmed by pyrosequencing. Of these, I would suggest high attention

on to 3 candidates (*EML2*, *HOXB1*, *RHOD*) which showed inverse correlation between DNA methylation and gene expression.

7.2.3. Further functional studies of *EFEMP1*

In other tumour types, multiple functions of fibulin-3 in tumourigenesis have been reported including tumour suppressor, apoptotic, angiogenic, metastatic, and invasion (Gallagher *et al.*, 2005; Obaya *et al.*, 2012; Seeliger *et al.*, 2009). My study showed for the first time fibulin-3 was significantly down-regulated in a large number of pituitary tumours in all 3 species investigated which are human, rat, and mouse. This phenomenon suggests that fibulin-3 may play roles in pituitary tumourigenesis. In my thesis, initial attempts had been done which shed some light toward the functions of this protein. For the subsequent studies I would suggest the following directions.

Enforced-expression of *EFEMP1* in other cells line model systems

Due to time limitation I had only been able to enforce the expression of *EFEMP1* in rat GH3 cells line which is corresponding to the GH subtype. In agreement with the literature reported for other tumour types, in case of pituitary tumours my data also suggest a bi-functional property of fibulin-3 where positive- and negative-correlations between this protein and *MMP2/MMP7* were observed in rat and human samples, respectively. Moreover, previous studies of pituitary tumourigenesis showed species- and/or subtype specificity of the target genes, for example the *BMP-4* behaved differently in lactotroph- and corticotroph-derived adenomas (Yacub-Usman *et al.*, 2012a). Therefore, it is necessary to enforce expression of *EFEMP1* in other pituitary cell lines corresponding to other species and tumour subtypes.

Investigate the effect of fibulin-3 on pituitary cell migration and invasion

My data suggest that fibulin-3 does not have effects neither on cell proliferation nor apoptosis pathway. However, the correlation between this protein and *MMP2/MMP7* suggest a function in cell migration and invasion which were observed in cervical and malignant cancers (Song *et al.*, 2011b; En-lin *et al.*, 2010; Hu *et al.*, 2012; Hu *et al.*, 2009). Therefore, it is apparent that cell invasion and migration assays should be performed on pituitary cell lines.

Section E: Appendices

Appendix 1: Primary pituitary adenomas used within the study

Discovery cohort

Study Number	Tumour Number	Classification	Grade
GH1	99	Somatotrophinoma	2
GH2	217	Somatotrophinoma	2
GH3	676	Somatotrophinoma	2
ACTH1	4	Corticotrophinoma	1
ACTH2	10	Corticotrophinoma	2
ACTH3	59	Corticotrophinoma	3
PRL1	25	Prolactinoma	2
PRL2	56	Prolactinoma	2
PRL3	57	Prolactinoma	1
NF1	551	Non-functioning	2
NF2	593	Non-functioning	2
NF3	598	Non-functioning	2

Investigation cohort

Study Number	Tumour Number	Classification	Grade
GH1	16	Somatotrophinoma	2
GH2	23	Somatotrophinoma	2
GH3	27	Somatotrophinoma	2
GH4	529	Somatotrophinoma	2
GH5	596	Somatotrophinoma	2
GH6	676	Somatotrophinoma	2
GH7	686	Somatotrophinoma	2
ACTH1	4	Corticotrophinoma	1
ACTH2	10	Corticotrophinoma	2
ACTH3	22	Corticotrophinoma	1
ACTH4	32	Corticotrophinoma	1
ACTH5	59	Corticotrophinoma	2
ACTH6	195	Corticotrophinoma	2
ACTH7	207	Corticotrophinoma	1
PRL1	25	Prolactinoma	2
PRL2	29	Prolactinoma	2
PRL3	46	Prolactinoma	1
PRL4	56	Prolactinoma	1
PRL5	57	Prolactinoma	1
PRL6	396	Prolactinoma	1
PRL7	535	Prolactinoma	1
NF1	512	Non-functioning	3
NF2	577	Non-functioning	2
NF3	579	Non-functioning	2
NF4	581	Non-functioning	2
NF5	594	Non-functioning	3
NF6	607	Non-functioning	3
NF7	608	Non-functioning	2
NF8	613	Non-functioning	3
NF9	618	Non-functioning	3
NF10	624	Non-functioning	2
NF11	684	Non-functioning	2
NF12	688	Non-functioning	2
NF13	689	Non-functioning	3

Appendix 2: Commonly used solutions

All reagents were purchased from Sigma-Aldrich unless stated otherwise.

2.1. DNA extraction

The reagents used for DNA extraction in addition to those in the materials and methods are as follows:

Lysis buffer: TE buffer (pH8.0), 0.5% SDS pH 7.2 and 200µg/mL of Proteinase K

TE buffer (pH 8.0)

10mM Tris HCl (pH 8.0)

1mM EDTA (pH 8.0)

Mix well and autoclave. The amount of TE buffer added to the reaction will vary between samples and between experiments. Cells require 1.5mL of TE buffer whereas tissue requires 3mL of TE buffer.

20% SDS

10g of SDS added to 50mL of H₂O and rotate on the rotary shaker at room temperature for 2 hours.

Proteinase K

Obtained from Sigma in a powdered form. Add H₂O to give a concentration of 10mg/mL. Add desired amount to give 200µg/mL final concentration in lysis buffer.

2.2. RNA extraction

The reagents stated below are used in the extraction of RNA in addition to those stated in the materials and methods:

Lysis buffer: 5M Guanidinium thiocyanate:

5M Guanidinium isothiocyanate

0.5% (w/v) Sarkosyl (N-lauroylsarcosine)

25mM sodium citrate

8% (v/v) β-mercaptoethanol

Mix well and autoclave

Chloroform Isoamyl-alcohol (24:1)

Add 239.6mL of chloroform to 10.4mL of isoamyl-alcohol (250ml total volume), mix well and store at 4°C

2M Sodium acetate pH4.0:

Add 82.03g of sodium acetate to 200mL of H₂O. Adjust the pH to 4.0 using glacial acetic acid, adjust the volume to 500mL by adding H₂O and autoclave.

70% Ethanol

Take 70mL of 100% ethanol and add 30mL of H₂O.

2.3. cDNA synthesis

All reagents arrived ready prepared (Promega, Southampton, UK). However the dNTP's require diluting.

Dilution of dNTP (Bioline, London, UK)

The dNTP's are purchased as a 100mM stock and, therefore, comprise 25mM of each of the four bases. For the majority of experiments the working concentration for each base was 5mM. Therefore 200µL of the supplied dNTP was added to 800µL of ddH₂O. Aliquots were stored at -20°C and thawed on no more than three occasions.

2.4. Gel electrophoresis

Products from RNA extraction, DNA extraction and PCR were in some cases resolved by gel electrophoresis. In these cases the following reagents were used:

TAE buffer (50X)

To 800ml ddH₂O add 242g Tris base, 57.1ml glacial acetic acid, 100ml 0.5M EDTA. Make up to 1X final volume using dH₂O.

10X Tris-Borate-EDTA (TBE) electrophoresis buffer (10X):

108g Tris base

55g Boric acid

40mL 0.5M EDTA (pH8.0)

Mix well and store at room temperature.

Loading buffer:

1mg/ml Bromophenol blue in 50% glycerol. Store aliquots at -20°C

Loading dye:

0.25% Bromophenol blue

0.25% Xylene cyanol FF

40% Sucrose.

Make up to required volume in dH₂O. Loading buffer is X6 therefore when loading 5µL of PCR product add 1µL of loading buffer. Mix well and load onto gel. If adding more or even less sample scale up or down the loading buffer quantities.

2.5. Bacterial transformation

DNA sequencing analysis:

DNA sequencing was performed commercially. In this case sequencing of native DNA or of DNA subject to sodium bisulphite conversion was performed. Converted DNA and amplified PCR product was first eluted from Gene Elute PCR column prior sub-cloning into the T:A vector prior to bacterial cloning and transformation (and as described in the M&M).

LB Medium:

Luria Broth is used for maintenance and propagation of E.coli. LB broth comprises Bacto-tryptone, Yeast and NaCl.

To 500mL of ddH₂O add 12.5g of LB broth mix well and autoclave immediately

LB agar plates containing ampicillin, IPTG and X-Gal:

Dissolve 12.5g of LB and 7.5g agar in 500mL of ddH₂O. Mix well and autoclave immediately. Cool to 50 °C before adding 60µL/mL ampicillin, 0.5mM IPTG, 0.08mg/ml X-gal. Mix solution avoiding frothing and pour into 90mm petri dishes at a depth of 3-5mm. Allow to set on the bench and further incubating at 37°C for complete setting. Once plates have set store at 4°C until ready to use.

Preparation of competent cells

Chemically Competent JM109 cells (Promega) were routinely prepared for transformation purposes. A colony from an LB plate was inoculated into 5mL of LB media and incubated for 16 hours at 37°C with shaking in the orbital shaker (150-200rpm). The entire 5mL culture was used to inoculate 250mL of LB medium containing 20mM MgSO₄ in a sterile glass flask and incubated at 37°C again with shaking. The culture was allowed to grow to an absorbance (OD₆₀₀) of 0.4-0.6 (typically 3-5 hours). The cells were collected by centrifugation at 914 x g for 10 minutes at 4°C using IEC Centra-8R Centrifuge (International equipment company, USA) a temperature controlled centrifuge. The supernatant was discarded and the cells were re-suspended in 10mL of ice cold chemical buffer 1: TFB1 (see below) and incubated on ice for 30 minutes. Cells were then centrifuged at 914 x g for 10 minutes at 4°C. The supernatant was removed again and cells were re-suspended in chemical buffer 2: TFB2 (see below) and incubated for a further 30 minutes on ice. This was then aliquoted into 0.5mL cold microcentrifuge tubes. These aliquoted chemically competent cells were then flash frozen by liquid nitrogen and were placed at -80°C until ready to use, and were used once without refreezing.

TFB1:

30mM Potassium acetate

10mM CaCl₂

50mM MnCl₂

100mM RbCl

15% glycerol

Adjust pH to 5.8 with 1M acetic acid, Filter sterilise and store at room temperature.

TFB2:

10mM MOPS (4-Morpholinepropanesulfonic acid) pH6.5

75mM CaCl₂

10mM RbCl

15% glycerol

Adjust to pH 6.5 with 1M KOH, filter sterilise and store at room temperature.

Appendix 3: Primer sequences

3.1. Bisulphite sequencing primers

Primers were directed toward sodium bisulphite modified DNA.

Human GLRX

Forward primer 1	TTATTTTGTTATTTTAATTTTAGTAA
Reverse primer 1	TCAAAAAAACTATATCCTTTAAA
Forward primer 2	AGTAAAAAGTTGTTTTGGGAGTG
Reverse primer 2	ATCATCCTTTCCCCAAAAAATTTATA
Amplified bisulphite converted sequence and CpGs interrogated	AGTAAAAAGTTGTTTTGGGAGTGGATTTTTTTTTTTTTTTGATCGTA GTATGTTGGGATTTTTTTACGTGGTAGTCGTGGTGTAGTTTTGGTT TTGGTTTCGTTTTTTGAAGATATTTGGTTTTATTATAAAATTTTTTGG GGAAAGGATGAT

Human MBD3

Forward primer	TGGAGAAATTGTTTGGTTTTA
Reverse primer	AAAACTCTATCTCCATCAATCA
Amplified bisulphite converted sequence and CpGs interrogated	TGGAGAAATTGTTTGGTTTAAAGATTTTTTGTGATATTGGGTTTAT TCGATTGGGATGTTTTTTTATTGAAGATTCGTTTTATTTTAAAGTT TTTTTAGTTGTAGGTTTAGGGAGTAGGATAGGTATATTTATTTATGAT TGATTGATTGATTGATGGAGATAGAGTTTT

3.2. Pyrosequencing primers

Primers were directed toward sodium bisulphite modified DNA. The table also shows the number of CpG interrogated by the primer combinations

Biotinylated primers are underlined

Gene/CpG		
Human cg08047457	Forward primer 1	ATTTGGGTGTAGGGATTGT
	Reverse primer 1	<u>AACCAAACCCAATATAAAAACCTCTTC</u>
	Forward primer 2	GGGAAGGAGTTGAGGAGAGT
	Reverse primer 2	CCACCCAATAAATAAACCAATATATTACT
	Sequencing primer	TGTAGGGATTGTGGG
	Number of CpGs	3
	Sequence to analyse	GTTYGAAGGYGGGTTGGGYG
Human cg20289949	Forward primer 1	<u>GGTTTGGGTGAAGGAGAAT</u>
	Reverse primer 1	AAACCCAAACAACCTATCTCCATCT
	Forward primer 2	GTGAGTTTGGGATTTTGGTTAA
	Reverse primer 2	CACCTATCATATCATCTCCT
	Sequencing primer	ACTAAACTCACATAAACTTATTAC
	Number of CpGs	1
	Sequence to analyse	AAACCRAAAA CTAAAAAAA CCCC
Human cg06197492	Forward primer 1	AGATGGGGAGTAATGGAATGTTT
	Reverse primer 1	<u>ATATAAACCCACCTTAACAAATACCTATAC</u>
	Forward primer 2	AGTTTAGGGTTTTTGTGAAGT
	Reverse primer 2	TCCCTTCTAAATTTAATTTACACTAAATCA
	Sequencing primer	GGTTTAGATGGAGGG
	Number of CpGs	2
	Sequence to analyse	YGGTYGGGTTTTGTATAGGTATTTGT
Human cg24739326	Forward primer 1	TTGGGGATTTTAGGATTTAGGT
	Reverse primer 1	<u>CCTCCTTCTATTTTATCAAATAATCTTC</u>
	Forward primer 2	GTAGGGGTTTGGGTGTGTTAA
	Reverse primer 2	CTAATCTCTAAACCTAACTTCC
	Sequencing primer	AGTAGGGTAGGAGGTA
	Number of CpGs	2
	Sequence to analyse	GTTTGAGAGTYGGTAGTTTYGTTTTGGAGT TTGAGTA
Human <i>BCL9L</i>	Forward primer 1	GGTTTGTGAGTGTTTTTTGTGATTAT
	Reverse primer 1	<u>CCCCAACTCTAACTCACACACACT</u>
	Forward primer 2	AGGTTTTTGTGTTTGTGTAAGTTTG
	Reverse primer 2	AAACTCAAAAACCCACAAACACC
	Sequencing primer	GGTTTTAGTTTTTTTATTTGTTTAA
	Number of CpGs	6
	Sequence to analyse	TGGGTTYGGY GTTTTTATTT AGTTTTYGGG TAGTTTTYGGT GTTTGYGYGT GTTTTAGGTT
Human <i>COL1A2</i>	Forward primer 1	GGTGAGGAGGATTGGATAGT
	Reverse primer 1	<u>CCTCCCCCCCCCTTCCAAA</u>
	Forward primer 2	AGTGTATTATTGTTGAAATAGTGGG
	Reverse primer 2	CCTAACCTAAAATACCTCCAAA
	Sequencing primer	GATTGGATAGTTTTTGTTTTG
	Number of CpGs	5

	Sequence to analyse	ATYGTYGGAG ATTTGTAAAT TTGTTTATG TYGGGGTTGT AGAGTATTTY GAYGTGTTTT ATAGTGTGTTT
Human	Forward primer 1	GGTTTGTAGGGTTAGAGAGAGAAT
<i>CXCL14</i>	Reverse primer 1	<u>AAAAAAACCTTCCCAACCC</u>
	Forward primer 2	TTTTTAGGTTTTTGAGGATTGA
	Reverse primer 2	ATAACTCACTAAATTTCTCAAT
	Sequencing primer	GGGTTTTTTAAAGGTAGAG
	Number of CpGs	7
	Sequence to analyse	YGGGAYGTTG GAAATGTGTG YGYGTTGTGG TATGGGTGTG TAAGTGTGYG AAGGYGGYGT GTTGTGTGAG
Human	Forward primer 1	GTTAAGGGTGGTAAGGGGGAGTAT
<i>D4ST1</i>	Reverse primer 1	<u>AACCACCCCTATCCTCTAAACTT</u>
	Forward primer 2	TGGAGTGGTTAAGATTGTAGGA
	Reverse primer 2	ACCCCTACCCCAAAAAACG
	Sequencing primer	TTTGTAGTTGGAAAGATTAGTTT
	Number of CpGs	5
	Sequence to analyse	TYGGTAYGTT TGGATTYGGA GAGAGYGTG TTYGTTTTT TGTTTTTAGA TT
Human	Forward primer 1	TTGTTGTTAAGAAAGGGGATTTTATTT
<i>EFEMP1</i>	Reverse primer 1	CCTTCTCACATATCTCAAACTTCT
	Forward primer 2	GGGGATTTAGATTGATAAGATAAAAGAGA
	Reverse primer 2	<u>CCCCAACACAACTCTCAAAATA</u>
	Sequencing primer	ATTTGGGATTTTTAAATAAAGTT
	Number of CpGs	2
	Sequence to analyse	GTGYGTAGAG AAAGGYGTGG AAATGTTATT TTGAGAGTTT GT
Human	Forward primer 1	GGTTTTAGGGGTGGTTGT
<i>EML2</i>	Reverse primer 1	<u>TCACCCAACCTCCAAACTACTCAT</u>
	Forward primer 2	ATTAGTAGAATTGGGGTTTTAATTTA
	Reverse primer 2	TAAACCTAAAATCTAAACCCC
	Sequencing primer	GGTTAGGGAGTTTTTTAGGGT
	Number of CpGs	6
	Sequence to analyse	YGGAAGYGGT AGAGTYGGT TYGGTTTTAT TTTYGGTTY GAAGTTT
Human	Forward primer 1	GGGATAGAAAGATTTTTTAAGGGAGTATT
<i>FEM1C</i>	Reverse primer 1	<u>CAACCTAAACCTCACATCACTAATATAC</u>
	Forward primer 2	ATTTTTTTTTTGGGGTTTTTGAGAT
	Reverse primer 2	taaaccaaactataacctccaaccc
	Sequencing primer	GTAAGTAGGTAAGTTTTTTTAGTAG
	Number of CpGs	10
	Sequence to analyse	TTTAGGTTTTY GTTTTTTTTT YGTGTTTTA TTTYGGTYGY GAYGTTTTYGT TTYGYGTTA GGTGTTGTT GGT
Human	Forward primer 1	GTTGAGGTTTTGAGAGGTTAGTAATATG
<i>FLJ46380</i>	Reverse primer 1	<u>CCCCACTCTCCATCCCATTTA</u>
	Forward primer 2	ATTATTAGGATATTAGTTTTATTTTAA
	Reverse primer 2	ACTACAAAACCAACTATCTTA
	Sequencing primer	GGAAAGTAGGTGGAGG
	Number of CpGs	7
	Sequence to analyse	GGYGTTYGTT GTTTTTTGGY GTTTTTAATT GGTGTYGAG GAAATGAGGY GYGATGATT
Human	Forward primer 1	AATGAAAAAGAGAGGAGGTTTAG
<i>GLRX</i>	Reverse primer 1	CAAAATACCCATTCTCCTTATCCC

	Forward primer 2	<u>AGTAAAAAGTTGTTTTGGGAGTG</u>
	Reverse primer 2	ATCATCCTTTCCCCAAAAATTTTATA
	Sequencing primer	AAAAATCCCAAACATACTAC
	Number of CpGs	1
	Sequence to analyse	RATCAAAAAA AAAAAAAAAA ATCCAC
Human	Forward primer 1	AGTTTAGGGTTTTTGTGAAGT
<i>H19</i>	Reverse primer 1	TCCCTTCTAAATTTAATTTACACTAAATCA
	Forward primer 2	AGATGGGGAGTAATGGAATGTTT
	Reverse primer 2	<u>ATATAAACCCACCTTAACAAATACCTATAC</u>
	Sequencing primer	GGTTTAGATGGAGGG
	Number of CpGs	2
	Sequence to analyse	YGGTYGGGTT TTGTATAGGT ATTTGT
Human	Forward primer 1	GTGAGTTTGGGATTTTGGTTAA
<i>HAAO</i>	Reverse primer 1	CACCTATCATATCATCTCCT
Assay1	Forward primer 2	<u>GGTTTGGGTGAAGGAGAAAT</u>
	Reverse primer 2	AAACCCAAACAACCTATCTCCATCT
	Sequencing primer	ACTAAACTCACATAAACTTATTAC
	Number of CpGs	1
	Sequence to analyse	AAACCRAAAA CTAAAAAAAA CCCC
Human	Forward primer 1	GGGTTTAGATTAAAGTTTATTAGGGGGTAA
<i>HAAO</i>	Reverse primer 1	<u>TTCACCCAAACCTCACT</u>
Assay2	Forward primer 2	TTAGTTAGGGTTTGGAGT
	Reverse primer 2	ACTCACATAAACTTATTACAA
	Sequencing primer	GTTTATTAGGGGGTAAAG
	Number of CpGs	7
	Sequence to analyse	TTYGYGGGTT TTAAGTTTTT ATTAYGGGGT YGYGGGGTT TGTTTYGTTT YG
Human	Forward primer 1	<u>TTGGTTTGGGAGAGATTATATGTGT</u>
<i>HOXB1</i>	Reverse primer 1	ACACCTCTTACCCTACA
	Forward primer 2	AGGAGTTTATTGAAGGTTTTGG
	Reverse primer 2	AAATTCATCCTATTATAATCCATAC
	Sequencing primer	CTTACCCTACAACCTT
	Number of CpGs	6
	Sequence to analyse	TCRACAATAT ATCACAACRC TAAAACTCRA ATCCATAACC TAACCCRCCTC RCCTAAACAA ACRCTTCCCT
Human	Forward primer 1	TGAAGAATGTATGGAGGAATTGTAT
<i>KCNE3</i>	Reverse primer 1	<u>AAATACCTCCCCCAACTCCTTAAC</u>
	Forward primer 2	ATATTTTGGATTTAAAGGTTTTATTG
	Reverse primer 2	CTCTAACTCTAAACTCCAC
	Sequencing primer	GTTTTATGTTGGGTTTGT
	Number of CpGs	6
	Sequence to analyse	TGAGGTTTATG YGGAGYGAGA GGTTTYGGT TTTTGGTTTA GGAAGTAGGY GTTGTAGGGG AYGAAGGGT YGG
Human	Forward primer 1	GTTGTATAGTTTTAGGAGGTAGGTT
<i>KIAA0676</i>	Reverse primer 1	<u>ACAACCCATTAAACCAACACCT</u>
	Forward primer 2	GTTTGGATTTAGGTCGGAGG
	Reverse primer 2	CGACGTCGCTACAACACGAA
	Sequencing primer	TTTAGGAGGTAGGTTG
	Number of CpGs	8
	Sequence to analyse	GYGYGYGTTY GAGGYGGGG TTGAGGGGGY GGYGYGGAAT T
Human	Forward primer 1	GAGAGTTTTGTGGGGAGGAAGTA

<i>KIAA1822</i>	Reverse primer 1	<u>ACAACCCCCATCCCTCC</u>
	Forward primer 2	GGAGGTGGAATAGAAGTTTA
	Reverse primer 2	ACACCCAAAACGCCAACAA
	Sequencing primer	ATTTGGAAAGAGGTTAAAAGAG
	Number of CpGs	8
	Sequence to analyse	GTYGTGTYGT ATTTTGYGTY GYGTTTTTTT TTTTGTAGTT TYGTYGAGTG TTTYGAT
Human	Forward primer 1	GGTTGGAAAGTAGGAGTTAGATTATAGT
<i>MAL</i>	Reverse primer 1	<u>ACCCAAAAACCACTAAACAAATACTACC</u>
	Forward primer 2	GATTAAAGTGGTTTTGTTTTGTT
	Reverse primer 2	ACAAACTCAAAAATAAAAAACAAATC
	Sequencing primer	GGGGTTTAGAAGAGGTTTA
	Number of CpGs	10
	Sequence to analyse	GGGYGGTGTT YGYGGYGT TY GGGTYGGGT TTTYGGGGY TGGGYGGG GYGG
Human	Forward primer 1	AGGGAAGTTGGGAAGGT
<i>MT1G</i>	Reverse primer 1	<u>ACCCAACAACCAACCACTATTTTTATAATC</u>
	Forward primer 2	TTTAGTTATTGATTAGAGATTGGG
	Reverse primer 2	AATTTAACATCCCAACACAAAAC
	Sequencing primer	ATTTAGAGGGTTGGTGTA
	Number of CpGs	6
	Sequence to analyse	AGGGYGGGY GGGYGTGTTG YGTTYGGTT YGTTTTTGA TTATAAAAT AGT
Human	Forward primer 1	GGTTGGGTAGAGTGGGATAG
<i>PDLIM4</i>	Reverse primer 1	<u>CCACCAACCAAAACCCCAAC</u>
	Forward primer 2	AATTAGGCGTATAGGGA
	Reverse primer 2	CAAACACCCGTAAATA
	Sequencing primer	GAGTGGGATAGGG
	Number of CpGs	11
	Sequence to analyse	YGGYGGGTY GGGGYGGGY GYGGYGTTA GGYGGYGY YGTTG
Human	Forward primer 1	GGGAGAAGATGTGTTAATGT
<i>PGLYRP1</i>	Reverse primer 1	<u>CCCCAACTCACTTCCTTCTCTAA</u>
	Forward primer 2	ATGGTATTAGGTTTGATTGGG
	Reverse primer 2	TTACACCCTACTAAACTTCCT
	Sequencing primer	GAGAGGAGAGATAAGTAG
	Number of CpGs	7
	Sequence to analyse	GATYGGGATT TYGGTTYGG GTTTTTTAGA AAYGGYGGG YGAAGYGTT AGGGTTTGT G
Human	Forward primer 1	GGAGGTTTTTTGGGATTTTAAGT
<i>RAMP1</i>	Reverse primer 1	<u>ACACAAAACCCCAACCATAATACACA</u>
	Forward primer 2	AGGGGAAGATTTTAGGTGGAGAT
	Reverse primer 2	CTACTCACCAAAAACAACCAAAAC
	Sequencing primer	TGGGAGTTTTAGGTTTGGTATTTA
	Number of CpGs	10
	Sequence to analyse	YGYGGGAGT YGTAGTGGGT YGGTYGYGY GGTYGYGTT TTTTATYGT TT
Human	Forward primer 1	GGAGAGGAAGGGTTTAGTT
<i>RHOD</i>	Reverse primer 1	<u>CCCACCAAAACCACTTAAC</u>
	Forward primer 2	TTTAGGGATTAGACGGTTTGA
	Reverse primer 2	CATCGACGAAAACCATCAACAA

	Sequencing primer	GAGGAAGGGTTTTAGTTTA
	Number of CpGs	9
	Sequence to analyse	GTYGTTYGGG TTTTTTATTT GGTTTTYGGG GYGTTYGGTT YGTATTTTTT TTYGTYGTTT YGTTTTTTTT ATATTTTTG
Human	Forward primer 1	GTTTAGTATTTTTATTTGTTTAGGGATT
<i>SAMD11</i>	Reverse primer 1	<u>CCCTTACTACCCTAACAAAATTT</u>
	Forward primer 2	GAGGATTTGGTTTTGGTTTTGGTTA
	Reverse primer 2	AATACTACCCTTCCAATCCTTAACC
	Sequencing primer	GGATTTTTAGTAGTTGTTGGATAA
	Number of CpGs	9
	Sequence to analyse	YGGYGAGTYG TGTATYGAGG TGGAGTYGG YGTTAATYGY YGTTGTTTT A
Human	Forward primer 1	<u>GGTTGTTTAGTGGGGTGAT</u>
<i>ST6GALNAC6</i>	Reverse primer 1	CCCCCAACCCCTCACAT
	Forward primer 2	AGGTTGTTGTATTTTGGTGGTTATT
	Reverse primer 2	CGAAACCTAACCCCACTAA
	Sequencing primer	CCCCACCCAAACCTAAAAATA
	Number of CpGs	9
	Sequence to analyse	A ACRAAACCT CRAAACRCR AATAACTCCR CCCACRAACR ACRCRAAC
Human	Forward primer 1	TTATTATAATTGGAGTGTATGGAGTAGG
<i>TFAP2E</i>	Reverse primer 1	<u>AACCCCAAACACTCACCATAAC</u>
	Forward primer 2	AGAATTGTTATAGTTTGTTTTTTTAATG
	Reverse primer 2	AAAATTAAAAATCCAACACTCAAAC
	Sequencing primer	GTTTAGAGGAGGAAGG
	Number of CpGs	7
	Sequence to analyse	GYGGGYGTTG GGTATTYGT GATYGATTTT TTTAAGTYG ATTAGTGTTY GTTYGTTTTG TTTTATGGA TT
Human	Forward primer 1	<u>GGGAAGTTAATGAGAAAATTAGGTTAGT</u>
<i>UBTD1</i>	Reverse primer 1	CCCCTAAACCTATCACTCACCTAC
	Forward primer 2	GGTATTTGGTTGGGTGGGG
	Reverse primer 2	GTGTTGGGAGTTTGTATTTTTTTT
	Sequencing primer	ACTCACCTACCTAAC
	Number of CpGs	11
	Sequence to analyse	CCCRCCCCAA ACCRACCCCT CTCCTTCCR CCRACRCR ATACCCRCRC CRCCRCACC ACCCRCCCA AATAC
Rat	Forward primer	GTGGTTGTTGTTGTTGTTAGAT
<i>Efemp1</i>	Reverse primer	<u>ATACAATCCCTACCCATACAATACTTCT</u>
	Sequencing primer	GTTAGTTTAGATTATTAGATTG
	Number of CpGs	6
	Sequence to analyse	T T TTTTTTG YGTATTTTTT YGTTGTYGGG GTTGATTGT YGYGYGGTTT TTYGTAGGTA GGAGTTTTTA GTTAAGGT
Mouse	Forward primer 1	TTTTTGTTGTTGTTGTTGTTAGT
<i>Efemp1</i>	Reverse primer 1	<u>CAACAAATCCACCTCCATCTT</u>
	Forward primer 2	TGGTTGGAAGTTTGGGTAGG
	Reverse primer 2	CACAACTAACAAATTATAATCCCAAA
	Sequencing primer	GGTTTTTTAGTAGGTAGGAGTTTAA
	Number of CpGs	4
	Sequence to analyse	AGTYGAGGTY GTAAGATTG YGTTATATTA GYGAGAGTTA TAT

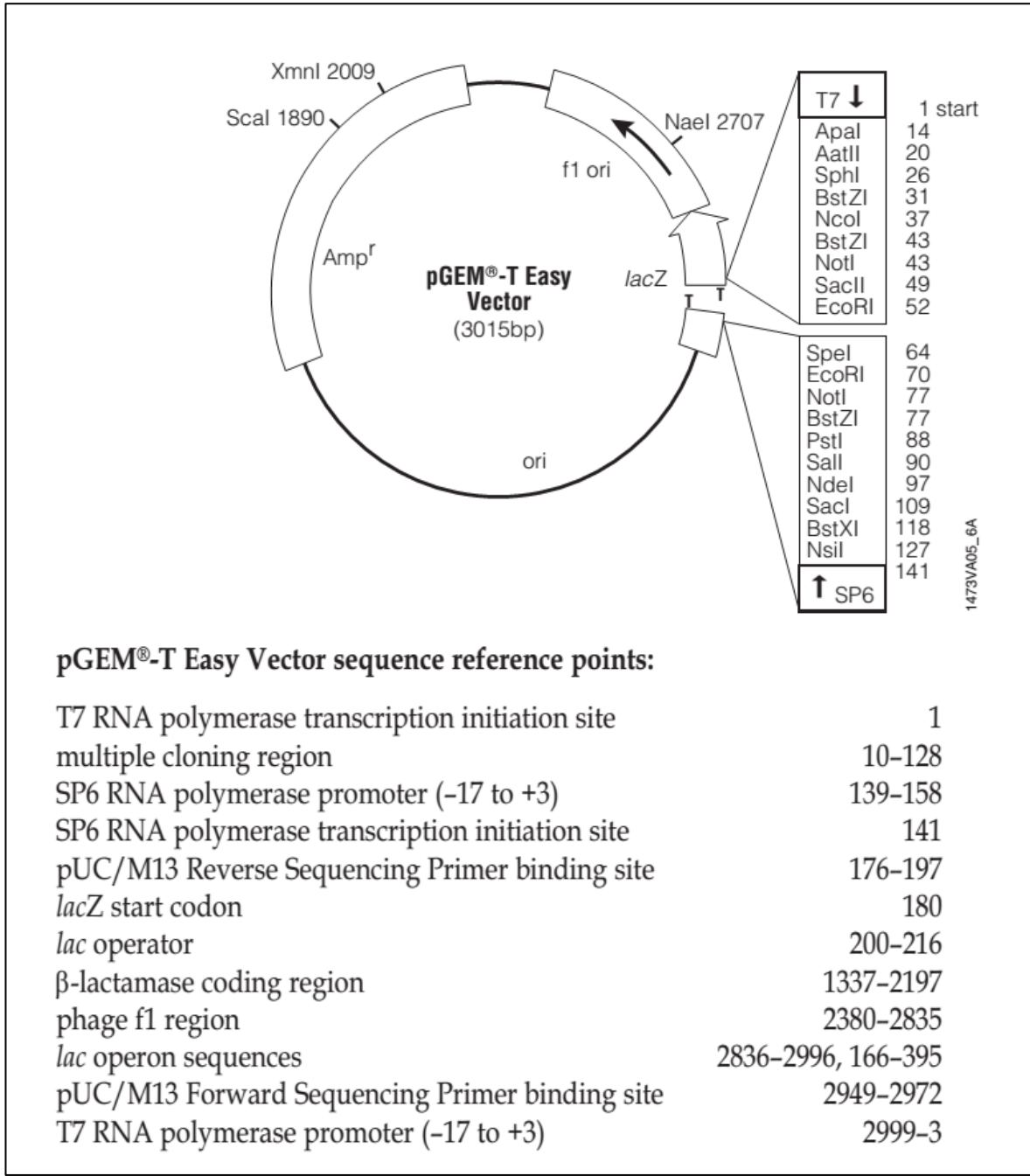
3.3. qRT-PCR primers

	Gene	Forward primer	Reverse primer	Amplicon size (bp)
Human	<i>GAPDH</i>	CGACCACTTTGTCAAGCTCA	GGGTCTTACTCCTTGGAGGC	102
	<i>BCL9L</i>	AGCAGGCTCAAGTCTCCTTC	GGGTACCCTGTTCCATGTTG	154
	<i>COL1A2</i>	TGTGGGTAGTCCTGGAGTCA	ACCAACAGGACCAGAAGGAC	232
	<i>EML2</i>	CTCATGATGGGAAGTTGCTG	ATGCTGTCATCCCACAAGAA	151
	<i>FLJ46380</i>	GCAGAGATGCAGAGCCTAAA	TCCCAAATCTCCTGGAAGAC	185
	<i>HAAO</i>	CCCATCATCCAGGAGTTCTT	CGATCACCTGGGTCTCATAG	211
	<i>HOXB1</i>	AAGACAGCGAAGGTGTCAGA	CGGCTCAGGTACTTGTGTGAA	119
	<i>KCNE3</i>	GCCAGCAGTCTGAGCTTCTA	CGCCTCTCTTCAGTCTGGTT	192
	<i>KIAA1822</i>	gccaacacctctctcaatgactt	gtacgggtagtagttgttgatgagg	261
	<i>MT1G</i>	ATGGACCCCAACTGCTCCTG	TCAGGCGCAGCAGCTGCAC	189
	<i>PGLYRP1</i>	GACGTGGGCTACAACCTCCT	CATGTAGTTGCCCATGAAGC	135
	<i>RHOD</i>	GCAGGGCAAGATGACTATGA	GGTAGGTCACAGGCTCCAAT	244
	<i>TFAP2E</i>	AATGTGACGCTGCTGACTTC	GGTCCTGAGCCATCAAGTCT	214
	<i>EFEMP1</i>	GTCACAGGACACCGAAGAAAC	TTGCATTGCTGTCTCACAGGA	81
	<i>MMP2</i>	TTGACGGTAAGGACGGACTC	ACTTGCACTACTCCCCATCG	153
	<i>MMP7</i>	GCTGACATCATGATTGGCTT	CCCATCAAATGGGTAGGAGT	60
	<i>MMP9</i>	CCGGACCAAGGATACAGTTT	CGGCACTGAGGAATGATCTA	81
	<i>VEGFA</i>	AGGGCAGAATCATCACGAAGT	AGGGTCTCGATTGGATGGCA	75
	<i>HIF1A</i>	GAACGTCGAAAAGAAAAGTCTCG	CCTTATCAAGATGCGAACTCACA	124
Rat	<i>Pbgd</i>	CAGCATCGCTACCACAGTGTC	ATGTCCGGTAACGGCGGC	126
	<i>Efemp1</i>	GGCCGAAATAACTTCGTCAT	CATCAATGTCTTGGCACACA	130
	<i>Mmp2</i>	GAAGGACAAGTGGTCCGAGT	TGCTGTATTCCCGACCATTA	88
	<i>Mmp7</i>	CCAAACAGTCCTAAGTGGCA	AAGAACCGAGGCAAGTCTGT	77
	<i>Mmp9</i>	GCTGGCAGAGGATTACCTGT	TGCTTCTCTCCCATCATCTG	66
	<i>Vegfa</i>	CAATGATGAAGCCCTGGAGT	TTTTTCGCGCTTTCGTTTTT	205
	<i>Hif1a</i>	GCAGCGATGACACGGAAA	CATATCGCTGTCCACATCAAA	342
Mouse	<i>Pbgd</i>	CGGGAAAACCCCTTGATGTC	TTCTTCTGGGTGCAAAATCTGG	285
	<i>Efemp1</i>	TGGCCGAAATAACTTTGTCA	TGCACTCATCAATATCTTGGC	137

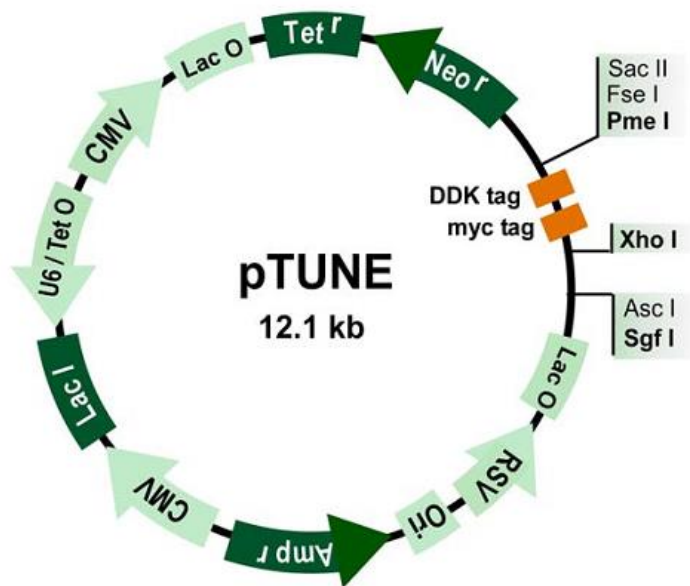
3.4. ChIP primers

Gene	Forward primer	Reverse primer	Amplicon size (bp)
Human <i>EFEMP1</i>	TGGTCTCTGAGGGAGATGCT	CCAGAGAGGACACGTTGACA	121
Rat <i>Efemp1</i>	TCGCAGAAGTGCATTGTACC	TCCATCTTCTCTGCCCAGTT	98
Mouse <i>Efemp1</i>	ACTGTCCCCGACTGTGAAAC	GTGATCCCAGACCAGCAAGT	62

Appendix 4: Vector maps



pGEM®-T Easy Vector Map and Sequence Reference Points



pTune_F
Kpn I
RBS
Kozac Consensus
Sgf I
Asc I

CCTCTAGAGTCGACCTGCAGCCGGAATTCGTCGACTGGATCCGGTACCGAGGAGATCTGCCGCCGCGATCGCCGGCGCG

Rsr II *
Mlu I *
Xho I
Myc Tag

CCAGATCTCAAGCTTAACTAGCTAGCGGACCG ACG CGT ACG CGG CCG CTC GAG CAG AAA CTC ATC TCA
 T R T R P L E Q K L I S

DDK Tag
Pme I

GAA GAG GAT CTG GCA GCA AAT GAT ATC CTG GAT TAC AAG GAT GAC GAC GAT AAG GTT TAA
 E E D L A A N D I L D Y K D D D D K V Stop

Fse I
Sac II
pTune_R

ACGGCCGGCCGGTCATAGCTGTTTCCTGAACAGATCCCTACACAAATCAGCGATTTT

* Do not use it as the cloning site.

Schematic of the pTUNE Inducible Vector and its Multiple Cloning Sites

Appendix 5: List of hypermethylated genes fulfil criteria of 0.4 in a single CpG site -

5.1. - within the CpG island

Gene	GH	NF	PRL	CUSH
SEPT9				
ACTA1				
ACTG1				
ACTN2				
ADAM11				
ADRA1A				
ADRA2C				
AGMAT				
ALX4				
APCDD1				
APOB				
AXIN1				
B3GNT5				
BAI1				
BCAR1				
BCL7B				
BCL9L				
BFSP1				
BIK				
BMP8A				
BTBD14A				
BZRP				
C10orf116				
C14orf58				
C16orf28				
C18orf22				
C19orf35				
C1orf59				
C1QTNF5				
C6orf150				
C7orf20				
C9orf112				
C9orf66				
CASP2				
CCDC69				

CD109				
CD9				
CDC42EP1				
CDCP1				
CDK5R1				
CHAT				
CHD5				
CHRNA1				
CHST8				
CLSTN1				
CNFN				
CNTN2				
COL1A2				
COL6A1				
CORO6				
CPXM2				
CRIP1				
CSMD2				
CTSZ				
CXCL14				
CYB561				
CYGB				
D4ST1				
DDAH2				
DHRS3				
DLX5				
DUSP2				
DYDC1				
EFEMP1				
EFNA2				
EFS				
ELN				
EMILIN3				
EML2				
ENG				
ENO3				
ENTPD2				
ERBB2				
ERBB3				
F3				
FAM69B				
FEM1C				
FLJ20032				
FLJ32569				
FLJ46380				

FLJ90166				
FMOD				
FOSL1				
FOXJ1				
FXYP7				
GABBR2				
GAPDHS				
GAS2L1				
GAS6				
GEFT				
GFRA3				
GNG12				
GPR133				
GSH2				
GSTM5				
HAAO				
HAS1				
HCLS1				
HCP5				
HDAC11				
HES6				
HEYL				
HOXB1				
HOXB2				
HSPA12B				
HTR6				
ICAM4				
IER3				
IFITM3				
IGF2BP1				
IGFBP1				
IL17R				
IQSEC1				
IRX2				
KAZALD1				
KBTD5				
KCNE3				
KCNH4				
KCNK4				
KCNQ1				
KCNQ1DN				
KIAA0676				
KIAA1446				
KIAA1822				
KIF17				

KLHL21				
KRT7				
LAMA5				
LDHC				
LHX3				
LOC153684				
LOC57228				
LYPD5				
MAL				
MCAM				
MEGF11				
MEIS3				
METRN				
MGC13168				
MGC14289				
MGC3207				
MGMT				
MMP14				
MPST				
MT1G				
MYL9				
NDRG2				
NEF3				
NEUROG1				
NODAL				
NPPB				
NPTX2				
NTF3				
NUDT4				
NUDT8				
OTOP3				
P4HA2				
PDLIM2				
PDLIM4				
PEX10				
PGAM2				
PGLYRP1				
PHLDA2				
PON3				
POU2F2				
POU3F1				
PPIE				
PPP1R3B				
PRKCDBP				
PSCD3				

PTPN5				
PYGM				
RAB11FIP4				
RAB34				
RAC2				
RAMP1				
RAP1GA1				
RAPGEFL1				
RASSF1				
RCN3				
RGL3				
RHOD				
RIPK3				
RTKN				
SAMD11				
SAMD4A				
SCARA3				
SEC14L4				
SECTM1				
SEMA3F				
SH2D3A				
SIPA1				
SLC15A3				
SLC25A10				
SLC2A4RG				
SLC5A1				
SLITL2				
SLMAP				
SNX8				
SOCS1				
SOCS2				
SPATS1				
ST6GALNAC6				
STK32C				
STUB1				
TBX4				
TBX6				
TCF15				
TCF21				
TCL1A				
TDRD5				
TFAP2E				
TINAGL1				
TLE1				
TMBIM1				

TMEM106A				
TMEM109				
TMEM41A				
TMEM44				
TMPRSS2				
TNFRSF10D				
TNFRSF13C				
TRIB2				
TRMU				
UAP1L1				
UBTD1				
UNQ9438				
VILL				
VWA1				
WFDC1				
WFIKKN2				
WNT4				
ZAR1				
ZNF135				
ZNF710				
Total	65	181	39	3

5.2. - outside of a CpG island

Gene	GH	NF	PRL	CUSH
ACADVL				
ACTN3				
ANXA9				
APOBEC2				
AXL				
Bles03				
BPI				
BST2				
C10orf26				
C1orf161				
C1orf38				
C1orf52				
C1QTNF3				
C20orf38				
CA9				
CARD15				
CD14				
CD248				
CD302				
CDC42EP5				
CDKN1A				
CFB				
CHI3L1				
CHRNA2				
CHRNA4				
CHRNE				
CHST13				
CKMT2				
CLCA1				
CLEC1A				
CLEC3A				
CLEC3B				
COL16A1				
COL5A2				
COL6A3				
COX6A2				
CX3CL1				
CYP4F12				
DCT				
DDR2				
EDAR				

EMP1				
EMP2				
FAM101A				
FAM107A				
FAM49A				
FBN3				
FBXL22				
FES				
FGF1				
FGFBP1				
FLJ20152				
FLJ20184				
FLJ20581				
FLJ25422				
FLJ34503				
FLJ36070				
FLJ43582				
FUT1				
FUT3				
GAL3ST4				
GDF5				
GIT1				
GLRX				
GPR92				
GRP				
GUCY2D				
HAPLN1				
HAPLN2				
HFE2				
HMGCS2				
HTR3A				
HYAL2				
IL16				
IL22RA1				
IL32				
INCA1				
INHBE				
ISLR				
KCNJ9				
KLK10				
KLK12				
LAMA4				
LGALS1				
LGI1				
LOC284948				

LOC340061				
LOXL3				
LRFN3				
LRRC8E				
LTC4S				
LYST				
MAP4K1				
MBD3				
MCEMP1				
MFAP4				
MGC18079				
MGC23244				
MGC4677				
MMP9				
MPG				
MRGPRF				
MRGPRX2				
MRVI1				
MT1H				
NFAM1				
NINJ2				
NRIP2				
NTF5				
NTN2L				
OASL				
PAQR4				
PAQR6				
PHKG1				
PLA2G3				
PLAT				
PLXNB1				
PRAP1				
PRKCG				
PRMT8				
PRTN3				
PRX				
PTMS				
RARRES3				
RASSF2				
RBP7				
RGS5				
S100A16				
S100A2				
SCGB3A2				
SDPR				

SERPINA12				
SERPING1				
SGCD				
SH3TC2				
SLC19A1				
SLC22A8				
SLC24A4				
SLC2A5				
SLC39A2				
SLC39A7				
SLC43A3				
SLC9A3R2				
SN				
SPON2				
STARD13				
STAT5A				
STIM1				
SULT1C1				
SUSD2				
TBX19				
TGM3				
TIMP3				
TIMP4				
TMEM71				
TMPRSS6				
TNFRSF14				
TNFSF15				
TRIP6				
TRPV6				
Total	32	145	10	1

Appendix 6: Gene ontology of hypermethylated genes

6.1. Go biological process

Term	Count	Genes
GO:0001501~skeletal system development	6	HOXB1, DLX5, COL1A2, ALX4, TCF15, BMP8A
GO:0007517~muscle organ development	5	ACTA1, ERBB2, ELN, ALX4, TCF15
GO:0007242~intracellular signaling cascade	11	SOCS2, RAC2, ERBB2, RASSF1, SOCS1, SIPA1, RAB34, COL1A2, EFS, ADRA1A, RHOD
GO:0032535~regulation of cellular component size	5	ACTA1, SIPA1, ELN, NPPB, VILL
GO:0007264~small GTPase mediated signal transduction	5	RAC2, RASSF1, RAB34, COL1A2, RHOD
GO:0007519~skeletal muscle tissue development	3	ACTA1, ERBB2, ELN
GO:0060538~skeletal muscle organ development	3	ACTA1, ERBB2, ELN
GO:0008015~blood circulation	4	ELN, COL1A2, NPPB, KCNQ1
GO:0003013~circulatory system process	4	ELN, COL1A2, NPPB, KCNQ1
GO:0048545~response to steroid hormone stimulus	4	SOCS2, ACTA1, ERBB2, SOCS1
GO:0007265~Ras protein signal transduction	3	RASSF1, COL1A2, RHOD
GO:0007010~cytoskeleton organization	5	ACTA1, RAC2, SIPA1, ELN, VILL
GO:0048705~skeletal system morphogenesis	3	HOXB1, ALX4, TCF15
GO:0014706~striated muscle tissue development	3	ACTA1, ERBB2, ELN
GO:0048483~autonomic nervous system development	2	HOXB1, ERBB2
GO:0060537~muscle tissue development	3	ACTA1, ERBB2, ELN
GO:0032570~response to progesterone stimulus	2	ERBB2, SOCS1
GO:0048562~embryonic organ morphogenesis	3	HOXB1, DLX5, ALX4
GO:0009952~anterior/posterior pattern formation	3	HOXB1, ALX4, TCF15
GO:0006936~muscle contraction	3	ACTA1, ADRA1A, KCNQ1
GO:0048741~skeletal muscle fiber development	2	ACTA1, ERBB2
GO:0043588~skin development	2	COL1A2, TCF15
GO:0040008~regulation of growth	4	SOCS2, SOCS1, SIPA1, NPPB
GO:0007167~enzyme linked receptor protein signaling pathway	4	WFIKK2, SOCS2, ERBB2, COL1A2
GO:0043062~extracellular structure organization	3	ERBB2, ELN, COL1A2
GO:0003012~muscle system process	3	ACTA1, ADRA1A, KCNQ1
GO:0048568~embryonic organ development	3	HOXB1, DLX5, ALX4
GO:0048747~muscle fiber development	2	ACTA1, ERBB2
GO:0007585~respiratory gaseous exchange	2	ELN, TCF15
GO:0009725~response to hormone stimulus	4	SOCS2, ACTA1, ERBB2, SOCS1
GO:0007243~protein kinase cascade	4	SOCS2, ERBB2, SOCS1, ADRA1A
GO:0007266~Rho protein signal transduction	2	COL1A2, RHOD
GO:0008544~epidermis development	3	CNFN, COL1A2, TCF15
GO:0007259~JAK-STAT cascade	2	SOCS2, SOCS1
GO:0050679~positive regulation of epithelial cell proliferation	2	ERBB2, DLX5
GO:0001558~regulation of cell growth	3	SOCS2, SIPA1, NPPB
GO:0003002~regionalization	3	HOXB1, ALX4, TCF15
GO:0007398~ectoderm development	3	CNFN, COL1A2, TCF15
GO:0009719~response to endogenous stimulus	4	SOCS2, ACTA1, ERBB2, SOCS1
GO:0008361~regulation of cell size	3	ACTA1, SIPA1, NPPB
GO:0007601~visual perception	3	EML2, C1QTNF5, EFEMP1
GO:0050953~sensory perception of light stimulus	3	EML2, C1QTNF5, EFEMP1
GO:0055002~striated muscle cell development	2	ACTA1, ERBB2

GO:0032355~response to estradiol stimulus	2	SOCS2, SOCS1
GO:0030833~regulation of actin filament polymerization	2	ELN, VILL
GO:0008283~cell proliferation	4	CRIP1, ERBB2, SIPA1, ELN
GO:0030203~glycosaminoglycan metabolic process	2	HAS1, PGLYRP1
GO:0030036~actin cytoskeleton organization	3	ACTA1, RAC2, ELN
GO:0055001~muscle cell development	2	ACTA1, ERBB2
GO:0048771~tissue remodeling	2	RAC2, ERBB2
GO:0007423~sensory organ development	3	ERBB2, DLX5, TCF15
GO:0048704~embryonic skeletal system morphogenesis	2	HOXB1, ALX4
GO:0007179~transforming growth factor beta receptor signaling pathway	2	WFIKKN2, COL1A2
GO:0008064~regulation of actin polymerization or depolymerization	2	ELN, VILL
GO:0030029~actin filament-based process	3	ACTA1, RAC2, ELN
GO:0045765~regulation of angiogenesis	2	ERBB2, NPPB
GO:0030832~regulation of actin filament length	2	ELN, VILL
GO:0010627~regulation of protein kinase cascade	3	SECTM1, ERBB2, SOCS1
GO:0006022~aminoglycan metabolic process	2	HAS1, PGLYRP1
GO:0032271~regulation of protein polymerization	2	ELN, VILL
GO:0050678~regulation of epithelial cell proliferation	2	ERBB2, DLX5
GO:0007015~actin filament organization	2	ACTA1, ELN
GO:0007389~pattern specification process	3	HOXB1, ALX4, TCF15
GO:0007498~mesoderm development	2	SECTM1, TCF15
GO:0042592~homeostatic process	5	RAC2, ERBB2, HAAO, NPPB, KCNQ1
GO:0048706~embryonic skeletal system development	2	HOXB1, ALX4
GO:0035113~embryonic appendage morphogenesis	2	DLX5, ALX4
GO:0030326~embryonic limb morphogenesis	2	DLX5, ALX4
GO:0051146~striated muscle cell differentiation	2	ACTA1, ERBB2
GO:0032956~regulation of actin cytoskeleton organization	2	ELN, VILL
GO:0043254~regulation of protein complex assembly	2	ELN, VILL
GO:0042981~regulation of apoptosis	5	SOCS2, TNFRSF10D, ERBB2, BIK, ALX4
GO:0032970~regulation of actin filament-based process	2	ELN, VILL
GO:0030308~negative regulation of cell growth	2	SIPA1, NPPB
GO:0048598~embryonic morphogenesis	3	HOXB1, DLX5, ALX4
GO:0043067~regulation of programmed cell death	5	SOCS2, TNFRSF10D, ERBB2, BIK, ALX4
GO:0010941~regulation of cell death	5	SOCS2, TNFRSF10D, ERBB2, BIK, ALX4
GO:0043583~ear development	2	DLX5, TCF15
GO:0007605~sensory perception of sound	2	EML2, KCNQ1
GO:0015672~monovalent inorganic cation transport	3	SLC5A1, KCNQ1, KCNE3
GO:0045792~negative regulation of cell size	2	SIPA1, NPPB
GO:0035108~limb morphogenesis	2	DLX5, ALX4
GO:0035107~appendage morphogenesis	2	DLX5, ALX4
GO:0008217~regulation of blood pressure	2	COL1A2, NPPB
GO:0007178~transmembrane receptor protein serine/threonine kinase signaling pathway	2	WFIKKN2, COL1A2
GO:0048736~appendage development	2	DLX5, ALX4
GO:0050954~sensory perception of mechanical stimulus	2	EML2, KCNQ1
GO:0060173~limb development	2	DLX5, ALX4

GO:0030198~extracellular matrix organization	2	ELN, COL1A2
GO:0043009~chordate embryonic development	3	HOXB1, ALX4, TCF15
GO:0043627~response to estrogen stimulus	2	SOCS2, SOCS1
GO:0009792~embryonic development ending in birth or egg hatching	3	HOXB1, ALX4, TCF15
GO:0060249~anatomical structure homeostasis	2	RAC2, HAAO
GO:0016051~carbohydrate biosynthetic process	2	HAS1, CHST8
GO:0016052~carbohydrate catabolic process	2	LDHC, PGLYRP1
GO:0045926~negative regulation of growth	2	SIPA1, NPPB
GO:0005976~polysaccharide metabolic process	2	HAS1, PGLYRP1
GO:0043066~negative regulation of apoptosis	3	SOCS2, TNFRSF10D, ERBB2
GO:0043069~negative regulation of programmed cell death	3	SOCS2, TNFRSF10D, ERBB2
GO:0060548~negative regulation of cell death	3	SOCS2, TNFRSF10D, ERBB2
GO:0042692~muscle cell differentiation	2	ACTA1, ERBB2
GO:0030030~cell projection organization	3	RAC2, ERBB2, DLX5
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	3	HOXB1, ALX4, TFAP2E
GO:0048511~rhythmic process	2	ERBB2, PGLYRP1
GO:0050877~neurological system process	6	EML2, C1QTNF5, ERBB2, EFEMP1, KCNQ1, TCF15
GO:0042391~regulation of membrane potential	2	ERBB2, KCNQ1
GO:0051493~regulation of cytoskeleton organization	2	ELN, VILL
GO:0030155~regulation of cell adhesion	2	ERBB2, SIPA1
GO:0030855~epithelial cell differentiation	2	DLX5, CNFN
GO:0032989~cellular component morphogenesis	3	ACTA1, ERBB2, DLX5
GO:0050878~regulation of body fluid levels	2	NPPB, ENTPD2
GO:0044087~regulation of cellular component biogenesis	2	ELN, VILL
GO:0050801~ion homeostasis	3	ERBB2, NPPB, KCNQ1
GO:0008284~positive regulation of cell proliferation	3	RAC2, ERBB2, DLX5
GO:0043434~response to peptide hormone stimulus	2	SOCS2, SOCS1
GO:0009101~glycoprotein biosynthetic process	2	ST6GALNAC6, CHST8
GO:0006813~potassium ion transport	2	KCNQ1, KCNE3
GO:0010033~response to organic substance	4	SOCS2, ACTA1, ERBB2, SOCS1
GO:0016265~death	4	TNFRSF10D, ADRA1A, BIK, TCF15
GO:0006357~regulation of transcription from RNA polymerase II promoter	4	HOXB1, ALX4, TFAP2E, TCF15
GO:0010740~positive regulation of protein kinase cascade	2	SECTM1, ERBB2
GO:0030001~metal ion transport	3	SLC5A1, KCNQ1, KCNE3
GO:0019725~cellular homeostasis	3	ERBB2, HAAO, KCNQ1
GO:0007610~behavior	3	RAC2, PGLYRP1, TCF15
GO:0040007~growth	2	ACTA1, BMP8A
GO:0045893~positive regulation of transcription, DNA-dependent	3	HOXB1, ALX4, TFAP2E
GO:0051254~positive regulation of RNA metabolic process	3	HOXB1, ALX4, TFAP2E
GO:0042060~wound healing	2	ERBB2, ENTPD2
GO:0042127~regulation of cell proliferation	4	RAC2, ERBB2, DLX5, ADRA1A
GO:0009617~response to bacterium	2	SOCS1, PGLYRP1
GO:0007409~axonogenesis	2	ERBB2, DLX5
GO:0009100~glycoprotein metabolic process	2	ST6GALNAC6, CHST8
GO:0007600~sensory perception	4	EML2, C1QTNF5, EFEMP1, KCNQ1
GO:0006916~anti-apoptosis	2	SOCS2, TNFRSF10D
GO:0048878~chemical homeostasis	3	ERBB2, NPPB, KCNQ1
GO:0048667~cell morphogenesis involved in neuron differentiation	2	ERBB2, DLX5

GO:0048812~neuron projection morphogenesis	2	ERBB2, DLX5
GO:0042493~response to drug	2	ERBB2, SOCS1
GO:0033043~regulation of organelle organization	2	ELN, VILL
GO:0009968~negative regulation of signal transduction	2	SOCS2, SOCS1
GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	2	SOCS2, ERBB2
GO:0060429~epithelium development	2	DLX5, CNFN
GO:0006812~cation transport	3	SLC5A1, KCNQ1, KCNE3
GO:0045941~positive regulation of transcription	3	HOXB1, ALX4, TFAP2E
GO:0000904~cell morphogenesis involved in differentiation	2	ERBB2, DLX5
GO:0048858~cell projection morphogenesis	2	ERBB2, DLX5
GO:0007166~cell surface receptor linked signal transduction	7	WFIKK2, SOCS2, ERBB2, COL1A2, NPPB, ADRA1A, ENTPD2
GO:0010648~negative regulation of cell communication	2	SOCS2, SOCS1
GO:0051056~regulation of small GTPase mediated signal transduction	2	ERBB2, SIPA1
GO:0010628~positive regulation of gene expression	3	HOXB1, ALX4, TFAP2E
GO:0050890~cognition	4	EML2, C1QTNF5, EFEMP1, KCNQ1
GO:0031175~neuron projection development	2	ERBB2, DLX5
GO:0032990~cell part morphogenesis	2	ERBB2, DLX5
GO:0006915~apoptosis	3	TNFRSF10D, ADRA1A, BIK
GO:0012501~programmed cell death	3	TNFRSF10D, ADRA1A, BIK
GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	3	HOXB1, ALX4, TFAP2E
GO:0009967~positive regulation of signal transduction	2	SECTM1, ERBB2
GO:0051173~positive regulation of nitrogen compound metabolic process	3	HOXB1, ALX4, TFAP2E
GO:0010557~positive regulation of macromolecule biosynthetic process	3	HOXB1, ALX4, TFAP2E
GO:0044057~regulation of system process	2	NPPB, KCNQ1
GO:0031328~positive regulation of cellular biosynthetic process	3	HOXB1, ALX4, TFAP2E
GO:0010647~positive regulation of cell communication	2	SECTM1, ERBB2
GO:0009891~positive regulation of biosynthetic process	3	HOXB1, ALX4, TFAP2E
GO:0048666~neuron development	2	ERBB2, DLX5
GO:0008219~cell death	3	TNFRSF10D, ADRA1A, BIK
GO:0000902~cell morphogenesis	2	ERBB2, DLX5
GO:0008285~negative regulation of cell proliferation	2	ERBB2, ADRA1A
GO:0006873~cellular ion homeostasis	2	ERBB2, KCNQ1
GO:0055082~cellular chemical homeostasis	2	ERBB2, KCNQ1
GO:0006811~ion transport	3	SLC5A1, KCNQ1, KCNE3
GO:0009057~macromolecule catabolic process	3	SOCS2, SOCS1, PGLYRP1
GO:0030182~neuron differentiation	2	ERBB2, DLX5
GO:0010604~positive regulation of macromolecule metabolic process	3	HOXB1, ALX4, TFAP2E
GO:0042325~regulation of phosphorylation	2	ERBB2, SOCS1
GO:0051174~regulation of phosphorus metabolic process	2	ERBB2, SOCS1
GO:0019220~regulation of phosphate metabolic process	2	ERBB2, SOCS1
GO:0032504~multicellular organism reproduction	2	ERBB2, BIK

GO:0048609~reproductive process in a multicellular organism	2	ERBB2, BIK
GO:0006350~transcription	6	HOXB1, BCL9L, HDAC11, ALX4, TFAP2E, TCF15
GO:0009611~response to wounding	2	ERBB2, ENTPD2
GO:0006355~regulation of transcription, DNA-dependent	5	HOXB1, DLX5, ALX4, TFAP2E, TCF15
GO:0055085~transmembrane transport	2	SLC5A1, KCNQ1
GO:0043632~modification-dependent macromolecule catabolic process	2	SOCS2, SOCS1
GO:0019941~modification-dependent protein catabolic process	2	SOCS2, SOCS1
GO:0051252~regulation of RNA metabolic process	5	HOXB1, DLX5, ALX4, TFAP2E, TCF15
GO:0051603~proteolysis involved in cellular protein catabolic process	2	SOCS2, SOCS1
GO:0006508~proteolysis	3	EML2, SOCS2, SOCS1
GO:0044257~cellular protein catabolic process	2	SOCS2, SOCS1
GO:0045449~regulation of transcription	7	HOXB1, DLX5, BCL9L, HDAC11, ALX4, TFAP2E, TCF15
GO:0030163~protein catabolic process	2	SOCS2, SOCS1
GO:0055114~oxidation reduction	2	LDHC, HAAO
GO:0006955~immune response	2	SECTM1, PGLYRP1
GO:0007155~cell adhesion	2	HAS1, EFS
GO:0022610~biological adhesion	2	HAS1, EFS
GO:0044265~cellular macromolecule catabolic process	2	SOCS2, SOCS1
GO:0010605~negative regulation of macromolecule metabolic process	2	SOCS1, KCNQ1
GO:0007049~cell cycle	2	RASSF1, SEPT9
GO:0007186~G-protein coupled receptor protein signaling pathway	2	ADRA1A, ENTPD2
GO:0050865~regulation of cell activation	1	ERBB2
GO:0014866~skeletal myofibril assembly	1	ACTA1
GO:0034655~nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	1	ENTPD2
GO:0050808~synapse organization	1	ERBB2
GO:0045453~bone resorption	1	RAC2
GO:0008406~gonad development	1	BIK
GO:0050892~intestinal absorption	1	SLC5A1
GO:0043266~regulation of potassium ion transport	1	KCNE3
GO:0014888~striated muscle adaptation	1	ACTA1
GO:0003014~renal system process	1	NPPB
GO:0001656~metanephros development	1	SLC5A1
GO:0006793~phosphorus metabolic process	1	ERBB2
GO:0019805~quinolinate biosynthetic process	1	HAAO
GO:0042312~regulation of vasodilation	1	NPPB
GO:0030240~muscle thin filament assembly	1	ACTA1
GO:0051046~regulation of secretion	1	KCNQ1
GO:0019228~regulation of action potential in neuron	1	ERBB2
GO:0050672~negative regulation of lymphocyte proliferation	1	ERBB2
GO:0006461~protein complex assembly	1	SEPTIN 09 (SEPT9)
GO:0035282~segmentation	1	TCF15
GO:0046425~regulation of JAK-STAT cascade	1	SOCS1
GO:0032944~regulation of mononuclear cell proliferation	1	ERBB2
GO:0009991~response to extracellular stimulus	1	ACTA1
GO:0022612~gland morphogenesis	1	ERBB2

GO:0046394~carboxylic acid biosynthetic process	1	HAAO
GO:0046579~positive regulation of Ras protein signal transduction	1	ERBB2
GO:0010741~negative regulation of protein kinase cascade	1	SOCS1
GO:0007162~negative regulation of cell adhesion	1	SIPA1
GO:0010043~response to zinc ion	1	HAAO
GO:0002792~negative regulation of peptide secretion	1	KCNQ1
GO:0051249~regulation of lymphocyte activation	1	ERBB2
GO:0042749~regulation of circadian sleep/wake cycle	1	PGLYRP1
GO:0051347~positive regulation of transferase activity	1	ERBB2
GO:0046903~secretion	1	NPPB
GO:0051186~cofactor metabolic process	1	HAAO
GO:0032268~regulation of cellular protein metabolic process	1	SOCS1
GO:0019320~hexose catabolic process	1	LDHC
GO:0019932~second-messenger-mediated signaling	1	ERBB2
GO:0006766~vitamin metabolic process	1	HAAO
GO:0046365~monosaccharide catabolic process	1	LDHC
GO:0006769~nicotinamide metabolic process	1	HAAO
GO:0016049~cell growth	1	ACTA1
GO:0030148~sphingolipid biosynthetic process	1	ST6GALNAC6
GO:0009181~purine ribonucleoside diphosphate catabolic process	1	ENTPD2
GO:0015749~monosaccharide transport	1	SLC5A1
GO:0044275~cellular carbohydrate catabolic process	1	LDHC
GO:0007631~feeding behavior	1	TCF15
GO:0021561~facial nerve development	1	HOXB1
GO:0044270~nitrogen compound catabolic process	1	ENTPD2
GO:0007586~digestion	1	SLC5A1
GO:0045580~regulation of T cell differentiation	1	ERBB2
GO:0003006~reproductive developmental process	1	BIK
GO:0046626~regulation of insulin receptor signaling pathway	1	SOCS1
GO:0043648~dicarboxylic acid metabolic process	1	HAAO
GO:0019674~NAD metabolic process	1	HAAO
GO:0046164~alcohol catabolic process	1	LDHC
GO:0042446~hormone biosynthetic process	1	CHST8
GO:0007589~body fluid secretion	1	NPPB
GO:0031100~organ regeneration	1	SOCS1
GO:0016569~covalent chromatin modification	1	HDAC11
GO:0007411~axon guidance	1	ERBB2
GO:0006767~water-soluble vitamin metabolic process	1	HAAO
GO:0035150~regulation of tube size	1	NPPB
GO:0006790~sulfur metabolic process	1	CHST8
GO:0046849~bone remodeling	1	RAC2
GO:0003016~respiratory system process	1	TCF15
GO:0006814~sodium ion transport	1	SLC5A1
GO:0009150~purine ribonucleotide metabolic process	1	ENTPD2
GO:0033088~negative regulation of immature T cell proliferation in the thymus	1	ERBB2
GO:0051494~negative regulation of cytoskeleton organization	1	VILL

GO:0009166~nucleotide catabolic process	1	ENTPD2
GO:0021546~rhombomere development	1	HOXB1
GO:0006917~induction of apoptosis	1	BIK
GO:0034097~response to cytokine stimulus	1	SOCS1
GO:0045860~positive regulation of protein kinase activity	1	ERBB2
GO:0043589~skin morphogenesis	1	COL1A2
GO:0006486~protein amino acid glycosylation	1	ST6GALNAC6
GO:0042476~odontogenesis	1	COL1A2
GO:0070663~regulation of leukocyte proliferation	1	ERBB2
GO:0019362~pyridine nucleotide metabolic process	1	HAAO
GO:0007528~neuromuscular junction development	1	ERBB2
GO:0019363~pyridine nucleotide biosynthetic process	1	HAAO
GO:0010629~negative regulation of gene expression	1	KCNQ1
GO:0042552~myelination	1	ERBB2
GO:0050884~neuromuscular process controlling posture	1	TCF15
GO:0009185~ribonucleoside diphosphate metabolic process	1	ENTPD2
GO:0006023~aminoglycan biosynthetic process	1	HAS1
GO:0021604~cranial nerve structural organization	1	HOXB1
GO:0016458~gene silencing	1	KCNQ1
GO:0050732~negative regulation of peptidyl-tyrosine phosphorylation	1	SOCS1
GO:0009137~purine nucleoside diphosphate catabolic process	1	ENTPD2
GO:0006468~protein amino acid phosphorylation	1	ERBB2
GO:0006935~chemotaxis	1	RAC2
GO:0043406~positive regulation of MAP kinase activity	1	ERBB2
GO:0051094~positive regulation of developmental process	1	DLX5
GO:0001822~kidney development	1	SLC5A1
GO:0045936~negative regulation of phosphate metabolic process	1	SOCS1
GO:0009191~ribonucleoside diphosphate catabolic process	1	ENTPD2
GO:0021545~cranial nerve development	1	HOXB1
GO:0080010~regulation of oxygen and reactive oxygen species metabolic process	1	RAC2
GO:0043503~skeletal muscle fiber adaptation	1	ACTA1
GO:0021570~rhombomere 4 development	1	HOXB1
GO:0006006~glucose metabolic process	1	LDHC
GO:0007507~heart development	1	ERBB2
GO:0051259~protein oligomerization	1	SEPTIN 09 (SEPT9)
GO:0014065~phosphoinositide 3-kinase cascade	1	ERBB2
GO:0046883~regulation of hormone secretion	1	KCNQ1
GO:0050795~regulation of behavior	1	PGLYRP1
GO:0001894~tissue homeostasis	1	RAC2
GO:0010035~response to inorganic substance	1	HAAO
GO:0042330~taxis	1	RAC2
GO:0045187~regulation of circadian sleep/wake cycle, sleep	1	PGLYRP1
GO:0048532~anatomical structure arrangement	1	HOXB1
GO:0003091~renal water homeostasis	1	NPPB
GO:0033081~regulation of T cell differentiation in the thymus	1	ERBB2

GO:0043114~regulation of vascular permeability	1	NPPB
GO:0070271~protein complex biogenesis	1	SEPTIN 09 (SEPT9)
GO:0050670~regulation of lymphocyte proliferation	1	ERBB2
GO:0009110~vitamin biosynthetic process	1	HAAO
GO:0009179~purine ribonucleoside diphosphate metabolic process	1	ENTPD2
GO:0042698~ovulation cycle	1	ERBB2
GO:0042532~negative regulation of tyrosine phosphorylation of STAT protein	1	SOCS1
GO:0048729~tissue morphogenesis	1	COL1A2
GO:0051726~regulation of cell cycle	1	SIPA1
GO:0006687~glycosphingolipid metabolic process	1	ST6GALNAC6
GO:0002237~response to molecule of bacterial origin	1	SOCS1
GO:0046395~carboxylic acid catabolic process	1	PON3
GO:0051338~regulation of transferase activity	1	ERBB2
GO:0007623~circadian rhythm	1	PGLYRP1
GO:0021548~pons development	1	HOXB1
GO:0016045~detection of bacterium	1	PGLYRP1
GO:0050817~coagulation	1	ENTPD2
GO:0034654~nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	1	HAAO
GO:0006096~glycolysis	1	LDHC
GO:0007422~peripheral nervous system development	1	ERBB2
GO:0010927~cellular component assembly involved in morphogenesis	1	ACTA1
GO:0021783~preganglionic parasympathetic nervous system development	1	HOXB1
GO:0009628~response to abiotic stimulus	1	ACTA1
GO:0033674~positive regulation of kinase activity	1	ERBB2
GO:0048730~epidermis morphogenesis	1	COL1A2
GO:0006664~glycolipid metabolic process	1	ST6GALNAC6
GO:0030834~regulation of actin filament depolymerization	1	VILL
GO:0030146~diuresis	1	NPPB
GO:0009132~nucleoside diphosphate metabolic process	1	ENTPD2
GO:0007272~ensheathment of neurons	1	ERBB2
GO:0040029~regulation of gene expression, epigenetic	1	KCNQ1
GO:0035116~embryonic hindlimb morphogenesis	1	ALX4
GO:0001503~ossification	1	BMP8A
GO:0000270~peptidoglycan metabolic process	1	PGLYRP1
GO:0048512~circadian behavior	1	PGLYRP1
GO:0033084~regulation of immature T cell proliferation in the thymus	1	ERBB2
GO:0043068~positive regulation of programmed cell death	1	BIK
GO:0006195~purine nucleotide catabolic process	1	ENTPD2
GO:0042516~regulation of tyrosine phosphorylation of Stat3 protein	1	SOCS1
GO:0001756~somitogenesis	1	TCF15
GO:0042063~gliogenesis	1	ERBB2
GO:0051129~negative regulation of cellular component organization	1	VILL
GO:0051017~actin filament bundle formation	1	ELN
GO:0043149~stress fiber formation	1	ELN
GO:0008643~carbohydrate transport	1	SLC5A1

GO:0030879~mammary gland development	1	ERBB2
GO:0006733~oxidoreduction coenzyme metabolic process	1	HAAO
GO:0048339~paraxial mesoderm development	1	TCF15
GO:0048485~sympathetic nervous system development	1	ERBB2
GO:0042518~negative regulation of tyrosine phosphorylation of Stat3 protein	1	SOCS1
GO:0031400~negative regulation of protein modification process	1	SOCS1
GO:0050866~negative regulation of cell activation	1	ERBB2
GO:0046777~protein amino acid autophosphorylation	1	ERBB2
GO:0051250~negative regulation of lymphocyte activation	1	ERBB2
GO:0042364~water-soluble vitamin biosynthetic process	1	HAAO
GO:0046700~heterocycle catabolic process	1	ENTPD2
GO:0019318~hexose metabolic process	1	LDHC
GO:0009253~peptidoglycan catabolic process	1	PGLYRP1
GO:0032269~negative regulation of cellular protein metabolic process	1	SOCS1
GO:0042326~negative regulation of phosphorylation	1	SOCS1
GO:0003018~vascular process in circulatory system	1	NPPB
GO:0048608~reproductive structure development	1	BIK
GO:0045184~establishment of protein localization	1	RAB34
GO:0016568~chromatin modification	1	HDAC11
GO:0035137~hindlimb morphogenesis	1	ALX4
GO:0001655~urogenital system development	1	SLC5A1
GO:0034656~nucleobase, nucleoside and nucleotide catabolic process	1	ENTPD2
GO:0022602~ovulation cycle process	1	ERBB2
GO:0046686~response to cadmium ion	1	HAAO
GO:0015758~glucose transport	1	SLC5A1
GO:0009595~detection of biotic stimulus	1	PGLYRP1
GO:0007596~blood coagulation	1	ENTPD2
GO:0045786~negative regulation of cell cycle	1	SIPA1
GO:0006688~glycosphingolipid biosynthetic process	1	ST6GALNAC6
GO:0030104~water homeostasis	1	NPPB
GO:0060021~palate development	1	ALX4
GO:0045597~positive regulation of cell differentiation	1	DLX5
GO:0035121~tail morphogenesis	1	TCF15
GO:0002694~regulation of leukocyte activation	1	ERBB2
GO:0042471~ear morphogenesis	1	DLX5
GO:0010310~regulation of hydrogen peroxide metabolic process	1	RAC2
GO:0043500~muscle adaptation	1	ACTA1
GO:0009261~ribonucleotide catabolic process	1	ENTPD2
GO:0022410~circadian sleep/wake cycle process	1	PGLYRP1
GO:0031333~negative regulation of protein complex assembly	1	VILL
GO:0060206~estrous cycle phase	1	ERBB2
GO:0007626~locomotory behavior	1	RAC2
GO:0050863~regulation of T cell activation	1	ERBB2
GO:0043933~macromolecular complex subunit organization	1	SEPTIN 09 (SEPT9)
GO:0016054~organic acid catabolic process	1	PON3
GO:0007276~gamete generation	1	BIK

GO:0001508~regulation of action potential	1	ERBB2
GO:0043122~regulation of I-kappaB kinase/NF-kappaB cascade	1	SECTM1
GO:0009820~alkaloid metabolic process	1	HAAO
GO:0010563~negative regulation of phosphorus metabolic process	1	SOCS1
GO:0051188~cofactor biosynthetic process	1	HAAO
GO:0051276~chromosome organization	1	HDAC11
GO:0007588~excretion	1	NPPB
GO:0046578~regulation of Ras protein signal transduction	1	ERBB2
GO:0070664~negative regulation of leukocyte proliferation	1	ERBB2
GO:0009612~response to mechanical stimulus	1	ACTA1
GO:0043603~cellular amide metabolic process	1	HAAO
GO:0065003~macromolecular complex assembly	1	SEPTIN 09 (SEPT9)
GO:0009259~ribonucleotide metabolic process	1	ENTPD2
GO:0050905~neuromuscular process	1	TCF15
GO:0014066~regulation of phosphoinositide 3-kinase cascade	1	ERBB2
GO:0009108~coenzyme biosynthetic process	1	HAAO
GO:0030837~negative regulation of actin filament polymerization	1	VILL
GO:0045087~innate immune response	1	PGLYRP1
GO:0016044~membrane organization	1	RAB34
GO:0021675~nerve development	1	HOXB1
GO:0030031~cell projection assembly	1	RAC2
GO:0045785~positive regulation of cell adhesion	1	ERBB2
GO:0008104~protein localization	1	RAB34
GO:0008584~male gonad development	1	BIK
GO:0046467~membrane lipid biosynthetic process	1	ST6GALNAC6
GO:0060263~regulation of respiratory burst	1	RAC2
GO:0008610~lipid biosynthetic process	1	ST6GALNAC6
GO:0021610~facial nerve morphogenesis	1	HOXB1
GO:0048839~inner ear development	1	DLX5
GO:0019221~cytokine-mediated signaling pathway	1	SOCS1
GO:0048486~parasympathetic nervous system development	1	HOXB1
GO:0030278~regulation of ossification	1	DLX5
GO:0050891~multicellular organismal water homeostasis	1	NPPB
GO:0048871~multicellular organismal homeostasis	1	RAC2
GO:0009247~glycolipid biosynthetic process	1	ST6GALNAC6
GO:0060443~mammary gland morphogenesis	1	ERBB2
GO:0006163~purine nucleotide metabolic process	1	ENTPD2
GO:0060056~mammary gland involution	1	ERBB2
GO:0048585~negative regulation of response to stimulus	1	SOCS1
GO:0006952~defense response	1	PGLYRP1
GO:0043410~positive regulation of MAPKKK cascade	1	ERBB2
GO:0006928~cell motion	1	ERBB2
GO:0030835~negative regulation of actin filament depolymerization	1	VILL
GO:0002683~negative regulation of immune system process	1	ERBB2
GO:0035136~forelimb morphogenesis	1	ALX4
GO:0051057~positive regulation of small GTPase	1	ERBB2

mediated signal transduction		
GO:0043413~biopolymer glycosylation	1	ST6GALNAC6
GO:0043242~negative regulation of protein complex disassembly	1	VILL
GO:0046426~negative regulation of JAK-STAT cascade	1	SOCS1
GO:0030902~hindbrain development	1	HOXB1
GO:0022037~metencephalon development	1	HOXB1
GO:0051291~protein heterooligomerization	1	SEPTIN 09 (SEPT9)
GO:0031424~keratinization	1	CNFN
GO:0042752~regulation of circadian rhythm	1	PGLYRP1
GO:0048232~male gamete generation	1	BIK
GO:0030199~collagen fibril organization	1	COL1A2
GO:0048644~muscle organ morphogenesis	1	TCF15
GO:0010324~membrane invagination	1	RAB34
GO:0007622~rhythmic behavior	1	PGLYRP1
GO:0006026~aminoglycan catabolic process	1	PGLYRP1
GO:0046661~male sex differentiation	1	BIK
GO:0006024~glycosaminoglycan biosynthetic process	1	HAS1
GO:0009913~epidermal cell differentiation	1	CNFN
GO:0060416~response to growth hormone stimulus	1	SOCS2
GO:0009581~detection of external stimulus	1	PGLYRP1
GO:0006476~protein amino acid deacetylation	1	HDAC11
GO:0022600~digestive system process	1	SLC5A1
GO:0046676~negative regulation of insulin secretion	1	KCNQ1
GO:0006007~glucose catabolic process	1	LDHC
GO:0043586~tongue development	1	ERBB2
GO:0012502~induction of programmed cell death	1	BIK
GO:0007267~cell-cell signaling	1	ADRA1A
GO:0046546~development of primary male sexual characteristics	1	BIK
GO:0001775~cell activation	1	ENTPD2
GO:0006796~phosphate metabolic process	1	ERBB2
GO:0000271~polysaccharide biosynthetic process	1	HAS1
GO:0043269~regulation of ion transport	1	KCNE3
GO:0045619~regulation of lymphocyte differentiation	1	ERBB2
GO:0043549~regulation of kinase activity	1	ERBB2
GO:0010038~response to metal ion	1	HAAO
GO:0030216~keratinocyte differentiation	1	CNFN
GO:0046874~quinolinate metabolic process	1	HAAO
GO:0010001~glial cell differentiation	1	ERBB2
GO:0006325~chromatin organization	1	HDAC11
GO:0016525~negative regulation of angiogenesis	1	NPPB
GO:0010817~regulation of hormone levels	1	CHST8
GO:0042445~hormone metabolic process	1	CHST8
GO:0042130~negative regulation of T cell proliferation	1	ERBB2
GO:0018108~peptidyl-tyrosine phosphorylation	1	ERBB2
GO:0030212~hyaluronan metabolic process	1	HAS1
GO:0043408~regulation of MAPKKK cascade	1	ERBB2
GO:0045669~positive regulation of osteoblast differentiation	1	DLX5
GO:0048732~gland development	1	ERBB2
GO:0001568~blood vessel development	1	COL1A2
GO:0030239~myofibril assembly	1	ACTA1

GO:0007283~spermatogenesis	1	BIK
GO:0035115~embryonic forelimb morphogenesis	1	ALX4
GO:0046888~negative regulation of hormone secretion	1	KCNQ1
GO:0005996~monosaccharide metabolic process	1	LDHC
GO:0007548~sex differentiation	1	BIK
GO:0006091~generation of precursor metabolites and energy	1	LDHC
GO:0050730~regulation of peptidyl-tyrosine phosphorylation	1	SOCS1
GO:0033087~negative regulation of immature T cell proliferation	1	ERBB2
GO:0060209~estrus	1	ERBB2
GO:0007568~aging	1	SOCS2
GO:0016570~histone modification	1	HDAC11
GO:0007599~hemostasis	1	ENTPD2
GO:0014068~positive regulation of phosphoinositide 3-kinase cascade	1	ERBB2
GO:0032272~negative regulation of protein polymerization	1	VILL
GO:0045667~regulation of osteoblast differentiation	1	DLX5
GO:0042755~eating behavior	1	TCF15
GO:0006732~coenzyme metabolic process	1	HAAO
GO:0021571~rhombomere 5 development	1	HOXB1
GO:0031099~regeneration	1	SOCS1
GO:0006027~glycosaminoglycan catabolic process	1	PGLYRP1
GO:0006512~ubiquitin cycle	1	SOCS1
GO:0006665~sphingolipid metabolic process	1	ST6GALNAC6
GO:0060396~growth hormone receptor signaling pathway	1	SOCS2
GO:0043244~regulation of protein complex disassembly	1	VILL
GO:0043085~positive regulation of catalytic activity	1	ERBB2
GO:0015031~protein transport	1	RAB34
GO:0042129~regulation of T cell proliferation	1	ERBB2
GO:0009154~purine ribonucleotide catabolic process	1	ENTPD2
GO:0050796~regulation of insulin secretion	1	KCNQ1
GO:0000272~polysaccharide catabolic process	1	PGLYRP1
GO:0019748~secondary metabolic process	1	HAAO
GO:0016575~histone deacetylation	1	HDAC11
GO:0051048~negative regulation of secretion	1	KCNQ1
GO:0006029~proteoglycan metabolic process	1	CHST8
GO:0019439~aromatic compound catabolic process	1	PON3
GO:0060348~bone development	1	BMP8A
GO:0032870~cellular response to hormone stimulus	1	SOCS2
GO:0016053~organic acid biosynthetic process	1	HAAO
GO:0043460~response to long exposure to lithium ion	1	ACTA1
GO:0048015~phosphoinositide-mediated signaling	1	ERBB2
GO:0008645~hexose transport	1	SLC5A1
GO:0021754~facial nucleus development	1	HOXB1
GO:0043405~regulation of MAP kinase activity	1	ERBB2
GO:0060341~regulation of cellular localization	1	KCNQ1
GO:0016192~vesicle-mediated transport	1	RAB34
GO:0019953~sexual reproduction	1	BIK
GO:0008016~regulation of heart contraction	1	KCNQ1
GO:0050880~regulation of blood vessel size	1	NPPB
GO:0051301~cell division	1	SEPTIN 09 (SEPT9)

GO:0006349~genetic imprinting	1	KCNQ1
GO:0043065~positive regulation of apoptosis	1	BIK
GO:0019226~transmission of nerve impulse	1	ERBB2
GO:0045444~fat cell differentiation	1	SOCS1
GO:0045137~development of primary sexual characteristics	1	BIK
GO:0051899~membrane depolarization	1	KCNQ1
GO:0008045~motor axon guidance	1	ERBB2
GO:0018212~peptidyl-tyrosine modification	1	ERBB2
GO:0030168~platelet activation	1	ENTPD2
GO:0007050~cell cycle arrest	1	RASSF1
GO:0050830~defense response to Gram-positive bacterium	1	PGLYRP1
GO:0030213~hyaluronan biosynthetic process	1	HAS1
GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB cascade	1	SECTM1
GO:0051216~cartilage development	1	BMP8A
GO:0051606~detection of stimulus	1	PGLYRP1
GO:0009134~nucleoside diphosphate catabolic process	1	ENTPD2
GO:0008366~axon ensheathment	1	ERBB2
GO:0002791~regulation of peptide secretion	1	KCNQ1
GO:0010639~negative regulation of organelle organization	1	VILL
GO:0042472~inner ear morphogenesis	1	DLX5
GO:0001889~liver development	1	ERBB2
GO:0032945~negative regulation of mononuclear cell proliferation	1	ERBB2
GO:0042509~regulation of tyrosine phosphorylation of STAT protein	1	SOCS1
GO:0051693~actin filament capping	1	VILL
GO:0048857~neural nucleus development	1	HOXB1
GO:0033083~regulation of immature T cell proliferation	1	ERBB2
GO:0045859~regulation of protein kinase activity	1	ERBB2
GO:0046627~negative regulation of insulin receptor signaling pathway	1	SOCS1
GO:0001932~regulation of protein amino acid phosphorylation	1	SOCS1
GO:0006939~smooth muscle contraction	1	ADRA1A
GO:0030147~natriuresis	1	NPPB
GO:0009165~nucleotide biosynthetic process	1	HAAO
GO:0006897~endocytosis	1	RAB34
GO:0010959~regulation of metal ion transport	1	KCNE3
GO:0034404~nucleobase, nucleoside and nucleotide biosynthetic process	1	HAAO
GO:0002695~negative regulation of leukocyte activation	1	ERBB2
GO:0021602~cranial nerve morphogenesis	1	HOXB1
GO:0016310~phosphorylation	1	ERBB2
GO:0050868~negative regulation of T cell activation	1	ERBB2
GO:0046496~nicotinamide nucleotide metabolic process	1	HAAO
GO:0001933~negative regulation of protein amino acid phosphorylation	1	SOCS1
GO:0001944~vasculature development	1	COL1A2
GO:0031399~regulation of protein modification process	1	SOCS1

GO:0031032~actomyosin structure organization	1	ACTA1
GO:0043501~skeletal muscle adaptation	1	ACTA1
GO:0044271~nitrogen compound biosynthetic process	1	HAAO
GO:0032496~response to lipopolysaccharide	1	SOCS1
GO:0051051~negative regulation of transport	1	KCNQ1
GO:0042742~defense response to bacterium	1	PGLYRP1
GO:0042745~circadian sleep/wake cycle	1	PGLYRP1
GO:0021612~facial nerve structural organization	1	HOXB1
GO:0009135~purine nucleoside diphosphate metabolic process	1	ENTPD2
GO:0022402~cell cycle process	1	RASSF1
GO:0070085~glycosylation	1	ST6GALNAC6
GO:0070050~neuron maintenance	1	HAAO
GO:0048678~response to axon injury	1	ERBB2
GO:0051248~negative regulation of protein metabolic process	1	SOCS1
GO:0006643~membrane lipid metabolic process	1	ST6GALNAC6
GO:0030166~proteoglycan biosynthetic process	1	CHST8
GO:0010942~positive regulation of cell death	1	BIK
GO:0044093~positive regulation of molecular function	1	ERBB2

6.2. Go molecular function

Count	Genes
4	RAC2, RAB34, RHOD, SEPT9
2	SOCS2, SOCS1
4	RAC2, RAB34, RHOD, SEPT9
4	RAC2, RAB34, RHOD, SEPT9
4	RAC2, RAB34, RHOD, SEPT9
5	ERBB2, RASSF1, COL1A2, TFAP2E, PON3
3	ERBB2, ALX4, TFAP2E
3	SLC5A1, KCNQ1, KCNE3
2	PGLYRP1, HDAC11
2	ACTA1, VILL
4	ERBB2, ALX4, TFAP2E, PON3
2	ELN, COL1A2
2	SOCS2, COL1A2
2	KCNQ1, KCNE3
2	ERBB2, COL1A2
4	ACTA1, ELN, COL1A2, VILL
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
5	HOXB1, DLX5, ALX4, TFAP2E, TCF15
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
2	SECTM1, BMP8A
2	KCNQ1, KCNE3
3	HOXB1, DLX5, ALX4
2	WFIKKN2, SOCS1
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
2	HOXB1, EFS
2	TFAP2E, PON3
6	ACTA1, RAC2, ERBB2, RAB34, RHOD, SEPT9
6	ACTA1, RAC2, ERBB2, RAB34, RHOD, SEPT9
5	HOXB1, DLX5, ALX4, TFAP2E, TCF15
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
6	ACTA1, RAC2, ERBB2, RAB34, RHOD, SEPT9
2	KCNQ1, KCNE3
2	KCNQ1, KCNE3
3	EFEMP1, ENTPD2, VILL
2	ACTA1, VILL
2	ERBB2, SOCS1
6	ACTA1, RAC2, ERBB2, RAB34, RHOD, SEPT9
11	CRIP1, SLC5A1, RASSF1, PDLIM4, EFEMP1, HAAO, KCNQ1, ENTPD2, KCNE3, MT1G, VILL
11	CRIP1, SLC5A1, RASSF1, PDLIM4, EFEMP1, HAAO, KCNQ1, ENTPD2, KCNE3, MT1G, VILL
11	CRIP1, SLC5A1, RASSF1, PDLIM4, EFEMP1, HAAO, KCNQ1, ENTPD2, KCNE3, MT1G, VILL
5	HOXB1, DLX5, ALX4, TFAP2E, TCF15
4	CRIP1, RASSF1, PDLIM4, MT1G

5	CRIP1, RASSF1, PDLIM4, HAAO, MT1G
2	ACTA1, ERBB2
2	ACTA1, ERBB2
2	ACTA1, ERBB2
2	ACTA1, ERBB2
2	ACTA1, ERBB2
1	HDAC11
1	HAAO
1	SOCS1
1	ADRA1A
1	ERBB2
1	WFIKKN2
1	SLC5A1
1	VILL
1	SOCS1
1	SOCS2
1	WFIKKN2
1	SIPA1
1	COL1A2
1	HAAO
1	ERBB2
1	SIPA1
1	ENTPD2
1	PON3
1	CHST8
1	SLC5A1
1	SOCS1
1	SLC5A1
1	PON3
1	ERBB2
1	ADRA1A
1	PGLYRP1
1	EML2
1	SOCS1
1	SIPA1
1	ERBB2
1	ERBB2
1	SIPA1
1	SLC5A1
1	HDAC11
1	NPPB
1	ELN
1	RASSF1
1	ADRA1A
1	PGLYRP1
1	PGLYRP1
1	SIPA1
1	SLC5A1
1	ERBB2
1	ACTA1
1	EML2
1	PGLYRP1

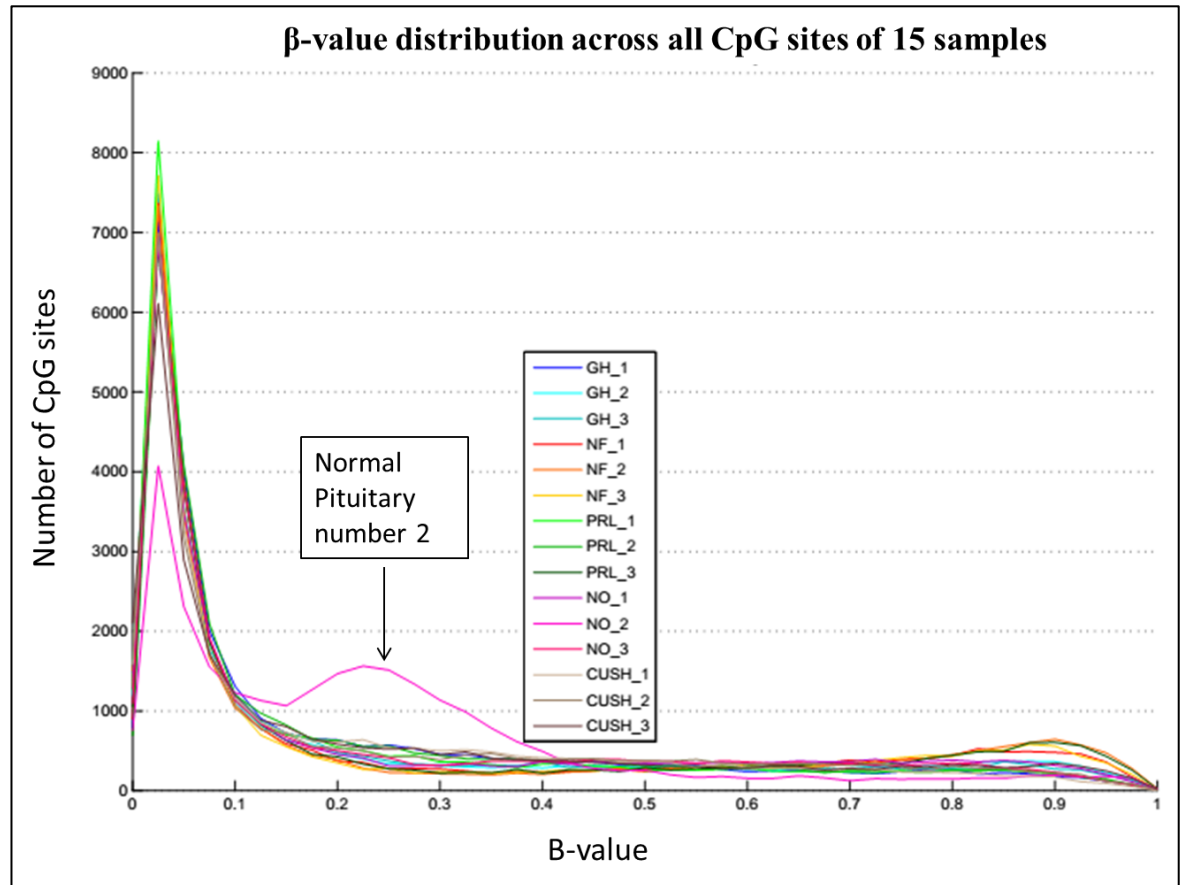
1	ST6GALNAC6
1	PGLYRP1
1	COL1A2
1	EML2
1	SOCS2
1	WFIKKN2
1	WFIKKN2
1	CHST8
1	SIPA1
1	SIPA1
1	PGLYRP1
1	SOCS2
1	TCF15
1	WFIKKN2
1	HAAO
1	EML2
1	HAS1
1	HDAC11
1	LDHC
1	RASSF1
1	SOCS2
1	SLC5A1
1	PON3
1	MT1G
1	PON3
1	SOCS1
1	KCNE3
1	PGLYRP1
1	ENTPD2
1	BMP8A
1	NPPB
1	PGLYRP1
1	HAAO
1	SOCS2
1	ERBB2
1	ERBB2
1	EML2
1	SLC5A1
1	HAAO
1	PGLYRP1
1	SOCS1
1	ERBB2
1	KCNQ1
1	ERBB2
1	EFS
1	EML2
1	SLC5A1
1	ERBB2
1	ACTA1
1	KCNQ1
1	LDHC
1	MT1G

1	CHST8
1	SOCS2
1	HAAO
1	ELN
1	WFIKKN2
1	SLC5A1
1	ERBB2
1	HDAC11
1	ERBB2
1	ADRA1A
1	HAAO
1	KCNE3

6.3. Kegg pathway

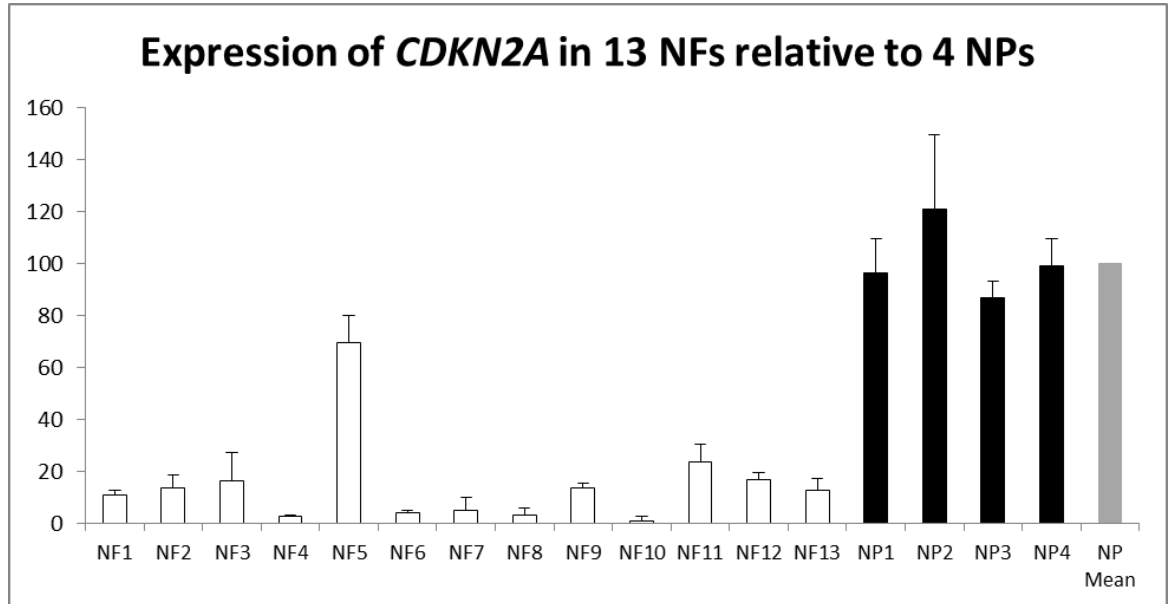
Term	Count	Genes
hsa04510:Focal adhesion	3	RAC2, ERBB2, COL1A2
hsa05219:Bladder cancer	2	ERBB2, RASSF1
hsa04930:Type II diabetes mellitus	2	SOCS2, SOCS1
hsa05223:Non-small cell lung cancer	2	ERBB2, RASSF1
hsa05212:Pancreatic cancer	2	RAC2, ERBB2
hsa04520:Adherens junction	2	RAC2, ERBB2
hsa05200:Pathways in cancer	3	RAC2, ERBB2, RASSF1
hsa04670:Leukocyte transendothelial migration	2	RAC2, SIPA1
hsa04360:Axon guidance	2	RAC2, RHOD
hsa04650:Natural killer cell mediated cytotoxicity	2	RAC2, TNFRSF10D
hsa04910:Insulin signaling pathway	2	SOCS2, SOCS1
hsa04630:Jak-STAT signaling pathway	2	SOCS2, SOCS1
hsa04020:Calcium signaling pathway	2	ERBB2, ADRA1A
hsa05110:Vibrio cholerae infection	1	KCNQ1
hsa00620:Pyruvate metabolism	1	LDHC
hsa04512:ECM-receptor interaction	1	COL1A2
hsa00230:Purine metabolism	1	ENTPD2
hsa05210:Colorectal cancer	1	RAC2
hsa04350:TGF-beta signaling pathway	1	BMP8A
hsa04666:Fc gamma R-mediated phagocytosis	1	RAC2
hsa04270:Vascular smooth muscle contraction	1	ADRA1A
hsa00604:Glycosphingolipid biosynthesis	1	ST6GALNAC6
hsa04120:Ubiquitin mediated proteolysis	1	SOCS1
hsa04062:Chemokine signaling pathway	1	RAC2
hsa04080:Neuroactive ligand-receptor interaction	1	ADRA1A
hsa04340:Hedgehog signaling pathway	1	BMP8A
hsa05213:Endometrial cancer	1	ERBB2
hsa00270:Cysteine and methionine metabolism	1	LDHC
hsa05215:Prostate cancer	1	ERBB2
hsa05416:Viral myocarditis	1	RAC2
hsa04012:ErbB signaling pathway	1	ERBB2
hsa04664:Fc epsilon RI signaling pathway	1	RAC2
hsa04370:VEGF signaling pathway	1	RAC2
hsa04310:Wnt signaling pathway	1	RAC2
hsa04662:B cell receptor signaling pathway	1	RAC2
hsa00640:Propanoate metabolism	1	LDHC
hsa04810:Regulation of actin cytoskeleton	1	RAC2
hsa04210:Apoptosis	1	TNFRSF10D
hsa00010:Glycolysis / Gluconeogenesis	1	LDHC
hsa04060:Cytokine-cytokine receptor interaction	1	TNFRSF10D
hsa00380:Tryptophan metabolism	1	HAAO
hsa04010:MAPK signaling pathway	1	RAC2

Appendix 7: Quality control of 15 pituitary samples in Illumina Infinium 27 BeadArray



A Kolmogorov–Smirnov test was used by our collaborator (Dr Richard Emes) to compare the distribution of β values between samples using the NIMBL Software. A single NP sample was identified as having a significantly different ($P < 0.05$) distribution to the other samples and was excluded from further analysis.

Appendix 8: Expression of *CDKN2A* in 13 NFs investigation cohort relative to 4 NPs



qRT-PCR determination of expression of *CDKN2A* at transcriptional level in 13 non-functioning pituitary tumours of the investigation cohort relative to mean of 4 normal pituitaries. White bars, Non-functioning tumours; Black bars, Normal pituitaries; Gray bars, Mean of 4 normal pituitaries; NF, Non-functioning; NP, Normal pituitary.

References

- Abbass, S. A., Asa, S. L. & Ezzat, S. (1997). Altered expression of fibroblast growth factor receptors in human pituitary adenomas. *J Clin Endocrinol Metab* 82(4): 1160-1166.
- Akiyama, Y., Maesawa, C., Ogasawara, S., Terashima, M. & Masuda, T. (2003). Cell-type-specific repression of the maspin gene is disrupted frequently by demethylation at the promoter region in gastric intestinal metaplasia and cancer cells. *Am J Pathol* 163(5): 1911-1919.
- Al-Azzawi, H., Yacqub-Usman, K., Richardson, A., Hofland, L. J., Clayton, R. N. & Farrell, W. E. (2011). Reversal of endogenous dopamine receptor silencing in pituitary cells augments receptor-mediated apoptosis. *Endocrinology* 152(2): 364-373.
- Albig, A. R., Neil, J. R. & Schiemann, W. P. (2006). Fibulins 3 and 5 Antagonize Tumor Angiogenesis In vivo. *Cancer Res* 66(5): 2621-2629.
- An, J. J., Cho, S. R., Jeong, D. W., Park, K. W., Ahn, Y. S. & Baik, J. H. (2003). Anti-proliferative effects and cell death mediated by two isoforms of dopamine D2 receptors in pituitary tumor cells. *Mol Cell Endocrinol* 206(1-2): 49-62.
- Asa, S. L. & Ezzat, S. (2002). The pathogenesis of pituitary tumours. *Nat Rev Cancer* 2(11): 836-849.
- Astuti, D., Latif, F., Wagner, K., Gentle, D., Cooper, W. N., Catchpoole, D., Grundy, R., Ferguson-Smith, A. C. & Maher, E. R. (2005). Epigenetic alteration at the DLK1-GTL2 imprinted domain in human neoplasia: analysis of neuroblastoma, pheochromocytoma and Wilms' tumour. *Br J Cancer* 92(8): 1574-1580.
- Bahar, A., Bicknell, J. E., Simpson, D. J., Clayton, R. N. & Farrell, W. E. (2004a). Loss of expression of the growth inhibitory gene GADD45gamma, in human pituitary adenomas, is associated with CpG island methylation. *Oncogene* 23(4): 936-944.
- Bahar, A., Simpson, D. J., Cutty, S. J., Bicknell, J. E., Hoban, P. R., Holley, S., Mourtada-Maarabouni, M., Williams, G. T., Clayton, R. N. & Farrell, W. E. (2004b). Isolation and characterization of a novel pituitary tumor apoptosis gene. *Mol Endocrinol* 18(7): 1827-1839.
- Ball, M. P., Li, J. B., Gao, Y., Lee, J. H., LeProust, E. M., Park, I. H., Xie, B., Daley, G. Q. & Church, G. M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* 27(4): 361-368.
- Ballestar, E., Paz, M. F., Valle, L., Wei, S., Fraga, M. F., Espada, J., Cigudosa, J. C., Huang, T. H. & Esteller, M. (2003). Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *EMBO J* 22(23): 6335-6345.
- Bates, A. S., Farrell, W. E., Bicknell, E. J., McNicol, A. M., Talbot, A. J., Broome, J. C., Perrett, C. W., Thakker, R. V. & Clayton, R. N. (1997). Allelic deletion in pituitary adenomas reflects aggressive biological activity and has potential value as a prognostic marker. *J Clin Endocrinol Metab* 82(3): 818-824.
- Belinsky, S. A., Klinge, D. M., Stidley, C. A., Issa, J. P., Herman, J. G., March, T. H. & Baylin, S. B. (2003). Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer. *Cancer Res* 63(21): 7089-7093.
- Berdasco, M. & Esteller, M. (2010). Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell* 19(5): 698-711.
- Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* 203(4): 971-983.
- Bibikova, M., Le, J., Barnes, B., Saedinia-Melnyk, S., Zhou, L., Shen, R. & Gunderson, K. L. (2009). Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics* 1(1): 177-200.
- Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E. W., Wu, B., Doucet, D., Thomas, N. J., Wang, Y., Vollmer, E., Goldmann, T., Seifart, C., Jiang, W., Barker, D. L., Chee, M. S., Floros, J. & Fan,

- J. B. (2006). High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 16(3): 383-393.
- Bilodeau, S., Vallette-Kasic, S., Gauthier, Y., Figarella-Branger, D., Brue, T., Berthelet, F., Lacroix, A., Batista, D., Stratakis, C., Hanson, J., Meij, B. & Drouin, J. (2006). Role of Brg1 and HDAC2 in GR trans-repression of the pituitary POMC gene and misexpression in Cushing disease. *Genes Dev* 20(20): 2871-2886.
- Bonazzi, V. F., Nancarrow, D. J., Stark, M. S., Moser, R. J., Boyle, G. M., Aoude, L. G., Schmidt, C. & Hayward, N. K. (2011). Cross-Platform Array Screening Identifies COL1A2, THBS1, TNFRSF10D and UCHL1 as Genes Frequently Silenced by Methylation in Melanoma. *PLoS ONE* 6(10): e26121.
- Brueckner, B., Stresemann, C., Kuner, R., Mund, C., Musch, T., Meister, M., Sultmann, H. & Lyko, F. (2007). The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. *Cancer Res* 67(4): 1419-1423.
- Byun, H.-M., Siegmund, K. D., Pan, F., Weisenberger, D. J., Kanel, G., Laird, P. W. & Yang, A. S. (2009). Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum Mol Genet* 18(24): 4808-4817.
- Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G. & Baylin, S. B. (1999). Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet* 21(1): 103-107.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. & Zhang, Y. (2002a). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298: 1039-1043.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. & Zhang, Y. (2002b). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298(5595): 1039-1043.
- Cheng, J. C., Matsen, C. B., Gonzales, F. A., Ye, W., Greer, S., Marquez, V. E., Jones, P. A. & Selker, E. U. (2003). Inhibition of DNA methylation and reactivation of silenced genes by zebularine. *J Natl Cancer Inst* 95(5): 399-409.
- Cheng, J. C., Weisenberger, D. J., Gonzales, F. A., Liang, G., Xu, G. L., Hu, Y. G., Marquez, V. E. & Jones, P. A. (2004a). Continuous zebularine treatment effectively sustains demethylation in human bladder cancer cells. *Mol Cell Biol* 24(3): 1270-1278.
- Cheng, J. C., Yoo, C. B., Weisenberger, D. J., Chuang, J., Wozniak, C., Liang, G., Marquez, V. E., Greer, S., Orntoft, T. F., Thykjaer, T. & Jones, P. A. (2004b). Preferential response of cancer cells to zebularine. *Cancer Cell* 6(2): 151-158.
- Chiba, T., Yokosuka, O., Arai, M., Tada, M., Fukai, K., Imazeki, F., Kato, M., Seki, N. & Saisho, H. (2004). Identification of genes up-regulated by histone deacetylase inhibition with cDNA microarray and exploration of epigenetic alterations on hepatoma cells. *J Hepatol* 41(3): 436-445.
- Chomczynski, P. & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162(1): 156-159.
- Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res* 22(15): 2990-2997.
- Cloos, P. A., Christensen, J., Agger, K. & Helin, K. (2008). Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev* 22(9): 1115-1140.
- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M. & Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452(7184): 215-219.
- Colella, S., Shen, L., Baggerly, K. A., Issa, J. P. & Krahe, R. (2003). Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 35(1): 146-150.

- Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L. K., Bartolomei, M. S., Rambow, F., Bassi, M. R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A. J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L. & Bellacosa, A. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146(1): 67-79.
- Cortez, C. C. & Jones, P. A. (2008). Chromatin, cancer and drug therapies. *Mutat Res* 647(1-2): 44-51.
- Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Su Huang, H., Petrelli, N. J., Zhang, X., O'Dorisio, M. S., Held, W. A., Cavenee, W. K. & Plass, C. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet* 24(2): 132-138.
- Costello, J. F. & Plass, C. (2001). Methylation matters. *J Med Genet* 38(5): 285-303.
- Cubas, P., Vincent, C. & Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401(6749): 157-161.
- Cummings, D. J., Tait, A. & Goddard, J. M. (1974). Methylated bases in DNA from *Paramecium aurelia*. *Biochim Biophys Acta* 374(1): 1-11.
- Daly, A. F., Burlacu, M. C., Livadariu, E. & Beckers, A. (2007). The epidemiology and management of pituitary incidentalomas. *Horm Res* 68 Suppl 5: 195-198.
- Dannenberg, L. O. & Edenberg, H. J. (2006). Epigenetics of gene expression in human hepatoma cells: expression profiling the response to inhibition of DNA methylation and histone deacetylation. *BMC Genomics* 7: 181.
- De Smet, C., De Backer, O., Faraoni, I., Lurquin, C., Brasseur, F. & Boon, T. (1996). The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. *Proc Natl Acad Sci U S A* 93(14): 7149-7153.
- de Vega, S., Iwamoto, T. & Yamada, Y. (2009). Fibulins: multiple roles in matrix structures and tissue functions. *Cell Mol Life Sci* 66(11-12): 1890-1902.
- Dean, F. B., Hosono, S., Fang, L., Wu, X., Faruqi, A. F., Bray-Ward, P., Sun, Z., Zong, Q., Du, Y., Du, J., Driscoll, M., Song, W., Kingsmore, S. F., Egholm, M. & Lasken, R. S. (2002). Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A* 99(8): 5261-5266.
- Deaton, A. M. & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development* 25(10): 1010-1022.
- Doerfler, W. (2006). The Almost-Forgotten Fifth Nucleotide in DNA: An Introduction. In *DNA Methylation: Basic Mechanisms*, Vol. 301, 3-18 (Eds W. Doerfler and P. Böhm). Springer Berlin Heidelberg.
- Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M. J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., Miller, J., Schlaeger, T., Daley, G. Q. & Feinberg, A. P. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* 41(12): 1350-1353.
- Dreos, R., Ambrosini, G., Cavin Périer, R. & Bucher, P. (2013). EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. *Nucleic Acids Research* 41(D1): D157-D164.
- Dudley, K. J., Revill, K., Clayton, R. N. & Farrell, W. E. (2009). Pituitary tumours: all silent on the epigenetics front. *Journal of Molecular Endocrinology* 42(6): 461-468.
- Dudley, K. J., Revill, K., Whitby, P., Clayton, R. N. & Farrell, W. E. (2008). Genome-wide analysis in a murine Dnmt1 knockdown model identifies epigenetically silenced genes in primary human pituitary tumors. *Mol Cancer Res* 6(10): 1567-1574.
- Duong, C. V., Emes, R. D., Wessely, F., Yacqub-Usman, K., Clayton, R. N. & Farrell, W. E. (2012). Quantitative, genome-wide analysis of the DNA methylome in sporadic pituitary adenomas. *Endocr Relat Cancer* 19(6): 805-816.

- Dupont, J.-M., Tost, J., Jammes, H. & Gut, I. G. (2004). De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Analytical Biochemistry* 333(1): 119-127.
- Dworakowska, D. & Grossman, A. B. (2009). The pathophysiology of pituitary adenomas. *Best Pract Res Clin Endocrinol Metab* 23(5): 525-541.
- Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., Burton, J., Cox, T. V., Davies, R., Down, T. A., Haefliger, C., Horton, R., Howe, K., Jackson, D. K., Kunde, J., Koenig, C., Liddle, J., Niblett, D., Otto, T., Pettett, R., Seemann, S., Thompson, C., West, T., Rogers, J., Olek, A., Berlin, K. & Beck, S. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 38(12): 1378-1385.
- Egger, G., Liang, G., Aparicio, A. & Jones, P. A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 429(6990): 457-463.
- Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. *Oncogene* 21(35): 5400-5413.
- Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics* 1(2): 239-259.
- Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A. & Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* 10(8): 2709-2721.
- Emes, R. D. & Farrell, W. E. (2012). Make way for the 'next generation': application and prospects for genome-wide, epigenome-specific technologies in endocrine research. *J Mol Endocrinol* 49(1): R19-27.
- En-lin, S., Sheng-guo, C. & Hua-qiao, W. (2010). The expression of EFEMP1 in cervical carcinoma and its relationship with prognosis. *Gynecol Oncol* 117(3): 417-422.
- Esteller, M. (2000). Epigenetic lesions causing genetic lesions in human cancer: promoter hypermethylation of DNA repair genes. *Eur J Cancer* 36(18): 2294-2300.
- Ezzat, S., Zhu, X., Loeper, S., Fischer, S. & Asa, S. L. (2006). Tumor-derived Ikaros 6 acetylates the Bcl-XL promoter to up-regulate a survival signal in pituitary cells. *Mol Endocrinol* 20(11): 2976-2986.
- Falls, J. G., Pulford, D. J., Wylie, A. A. & Jirtle, R. L. (1999). Genomic imprinting: implications for human disease. *Am J Pathol* 154(3): 635-647.
- Fang, J., Zhu, S., Xiao, S., Shi, Y., Jiang, S., Zhou, X. & Qian, L. (1996). Alterations of level of total genomic DNA methylation and pattern of c-myc, c-Ha-ras oncogene methylation in human gastric carcinogenesis. *Chin Med J (Engl)* 109(10): 787-791.
- Farrell, W. E. (2006). Pituitary tumours: findings from whole genome analyses. *Endocr Relat Cancer* 13(3): 707-716.
- Farrell, W. E. & Clayton, R. N. (2000). Molecular pathogenesis of pituitary tumors. *Front Neuroendocrinol* 21(3): 174-198.
- Fedele, M., Pierantoni, G. M., Visone, R. & Fusco, A. (2006). Critical role of the HMGA2 gene in pituitary adenomas. *Cell Cycle* 5(18): 2045-2048.
- Feil, R. & Berger, F. (2007). Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet* 23(4): 192-199.
- Feinberg, A. P., Ohlsson, R. & Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 7(1): 21-33.
- Fenaux, P., Mufti, G. J., Hellstrom-Lindberg, E., Santini, V., Finelli, C., Giagounidis, A., Schoch, R., Gattermann, N., Sanz, G., List, A., Gore, S. D., Seymour, J. F., Bennett, J. M., Byrd, J., Backstrom, J., Zimmerman, L., McKenzie, D., Beach, C. & Silverman, L. R. (2009). Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol* 10(3): 223-232.
- Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A., Marques, C. J., Andrews, S. & Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473(7347): 398-402.

- Flotho, C., Claus, R., Batz, C., Schneider, M., Sandrock, I., Ihde, S., Plass, C., Niemeyer, C. M. & Lubbert, M. (2009). The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells. *Leukemia* 23(6): 1019-1028.
- Fraga, M. F. & Esteller, M. (2005). Towards the human cancer epigenome: a first draft of histone modifications. *Cell Cycle* 4(10): 1377-1381.
- Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. & Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America* 89(5): 1827-1831.
- Frost, S. J., Simpson, D. J., Clayton, R. N. & Farrell, W. E. (1999). Transfection of an inducible p16/CDKN2A construct mediates reversible growth inhibition and G1 arrest in the AtT20 pituitary tumor cell line. *Mol Endocrinol* 13(11): 1801-1810.
- Fryer, A. A., Emes, R. D., Ismail, K. M., Haworth, K. E., Mein, C., Carroll, W. D. & Farrell, W. E. (2011). Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. *Epigenetics* 6(1): 86-94.
- Fuks, F., Burgers, W. A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet* 24(1): 88-91.
- Gallagher, W. M., Currid, C. A. & Whelan, L. C. (2005). Fibulins and cancer: friend or foe? *Trends Mol Med* 11(7): 336-340.
- Gardiner-Garden, M. & Frommer, M. (1987). CpG islands in vertebrate genomes. *J Mol Biol* 196(2): 261-282.
- Garinis, G. A., Patrinos, G. P., Spanakis, N. E. & Menounos, P. G. (2002). DNA hypermethylation: when tumour suppressor genes go silent. *Hum Genet* 111(2): 115-127.
- Gartler, S. M., Dyer, K. A. & Goldman, M. A. (1992). Mammalian X chromosome inactivation. *Mol Genet Med* 2: 121-160.
- Gebhard, C., Schwarzfischer, L., Pham, T. H., Schilling, E., Klug, M., Andreesen, R. & Rehli, M. (2006). Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia. *Cancer Res* 66(12): 6118-6128.
- Genereux, D. P., Johnson, W. C., Burden, A. F., Stoger, R. & Laird, C. D. (2008). Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. *Nucleic Acids Res* 36(22): e150.
- Giacomini, D., Paez-Pereda, M., Theodoropoulou, M., Labeur, M., Refojo, D., Gerez, J., Chervin, A., Berner, S., Losa, M., Buchfelder, M., Renner, U., Stalla, G. K. & Arzt, E. (2006). Bone morphogenetic protein-4 inhibits corticotroph tumor cells: involvement in the retinoic acid inhibitory action. *Endocrinology* 147(1): 247-256.
- Gil, E. Y., Jo, U. H., Jeong, H., Whang, Y. M., Woo, O. H., Cho, K. R., Seo, J. H., Kim, A., Lee, E. S., Koh, I., Kim, Y. H. & Park, K. H. (2012). Promoter methylation of RASSF1A modulates the effect of the microtubule-targeting agent docetaxel in breast cancer. *Int J Oncol* 41(2): 611-620.
- Gius, D., Cui, H., Bradbury, C. M., Cook, J., Smart, D. K., Zhao, S., Young, L., Brandenburg, S. A., Hu, Y., Bisht, K. S., Ho, A. S., Mattson, D., Sun, L., Munson, P. J., Chuang, E. Y., Mitchell, J. B. & Feinberg, A. P. (2004). Distinct effects on gene expression of chemical and genetic manipulation of the cancer epigenome revealed by a multimodality approach. *Cancer Cell* 6(4): 361-371.
- Gore, S. D., Baylin, S., Sugar, E., Carraway, H., Miller, C. B., Carducci, M., Grever, M., Galm, O., Dausies, T., Karp, J. E., Rudek, M. A., Zhao, M., Smith, B. D., Manning, J., Jiemjit, A., Dover, G., Mays, A., Zwiebel, J., Murgo, A., Weng, L. J. & Herman, J. G. (2006). Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. *Cancer Res* 66(12): 6361-6369.

- Gore, S. D. & Hermes-DeSantis, E. R. (2008). Future directions in myelodysplastic syndrome: newer agents and the role of combination approaches. *Cancer Control* 15 Suppl: 40-49.
- Gorovsky, M. A., Hattman, S. & Pleger, G. L. (1973). (6 N)methyl adenine in the nuclear DNA of a eucaryote, *Tetrahymena pyriformis*. *J Cell Biol* 56(3): 697-701.
- Greger, V., Passarge, E., Hopping, W., Messmer, E. & Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet* 83(2): 155-158.
- Grunau, C., Clark, S. J. & Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* 29(13): E65-65.
- Guo, Junjie U., Su, Y., Zhong, C., Ming, G.-I. & Song, H. (2011). Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain. *Cell* 145(3): 423-434.
- Ha, Y. S., Jeong, P., Kim, J. S., Kwon, W. A., Kim, I. Y., Yun, S. J., Kim, G. Y., Choi, Y. H., Moon, S. K. & Kim, W. J. (2012). Tumorigenic and prognostic significance of RASSF1A expression in low-grade (WHO grade 1 and grade 2) nonmuscle-invasive bladder cancer. *Urology* 79(6): 1411 e1411-1416.
- Han, H., Cortez, C. C., Yang, X., Nichols, P. W., Jones, P. A. & Liang, G. (2011). DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. *Hum Mol Genet* 20(22): 4299-4310.
- Hatada, I., Hayashizaki, Y., Hirotsune, S., Komatsubara, H. & Mukai, T. (1991). A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc Natl Acad Sci U S A* 88(21): 9523-9527.
- Hattman, S., Kenny, C., Berger, L. & Pratt, K. (1978). Comparative study of DNA methylation in three unicellular eucaryotes. *J Bacteriol* 135(3): 1156-1157.
- Hayward, B. E., Barlier, A., Korbonits, M., Grossman, A. B., Jacquet, P., Enjalbert, A. & Bonthron, D. T. (2001). Imprinting of the G(s)alpha gene *GNAS1* in the pathogenesis of acromegaly. *J Clin Invest* 107(6): R31-36.
- He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C. & Xu, G. L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333(6047): 1303-1307.
- Hebbes, T. R., Thorne, A. W. & Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J* 7(5): 1395-1402.
- Herman, J. G., Jen, J., Merlo, A. & Baylin, S. B. (1996). Hypermethylation-associated inactivation indicates a tumor suppressor role for p15INK4B. *Cancer Res* 56(4): 722-727.
- Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarr, J. R., Linehan, W. M. & et al. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A* 91(21): 9700-9704.
- Hodges, E., Smith, A. D., Kendall, J., Xuan, Z., Ravi, K., Rooks, M., Zhang, M. Q., Ye, K., Bhattacharjee, A., Brizuela, L., McCombie, W. R., Wigler, M., Hannon, G. J. & Hicks, J. B. (2009). High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. *Genome Res* 19(9): 1593-1605.
- Hoffmann, M. J. & Schulz, W. A. (2005). Causes and consequences of DNA hypomethylation in human cancer. *Biochem Cell Biol* 83(3): 296-321.
- Holliday, R. & Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. *Science* 187(4173): 226-232.
- Hotchkiss, R. D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem* 175(1): 315-332.
- Hu, B., Nandhu, M. S., Sim, H., Agudelo-Garcia, P. A., Saldivar, J. C., Dolan, C. E., Mora, M. E., Nuovo, G. J., Cole, S. E. & Viapiano, M. S. (2012). Fibulin-3 promotes glioma growth and resistance through a novel paracrine regulation of Notch signaling. *Cancer Res* 72(15): 3873-3885.

- Hu, B., Thirumara-Rajamani, K. K., Sim, H. & Viapiano, M. S. (2009). Fibulin-3 is uniquely upregulated in malignant gliomas and promotes tumor cell motility and invasion. *Mol Cancer Res* 7(11): 1756-1770.
- Hu, Y., Pioli, P. D., Siegel, E., Zhang, Q., Nelson, J., Chaturvedi, A., Mathews, M. S., Ro, D. I., Alkafeef, S., Hsu, N., Hamamura, M., Yu, L., Hess, K. R., Tromberg, B. J., Linskey, M. E. & Zhou, Y. H. (2011). EFEMP1 suppresses malignant glioma growth and exerts its action within the tumor extracellular compartment. *Mol Cancer* 10: 123.
- Huang, T. H.-M., Perry, M. R. & Laux, D. E. (1999). Methylation Profiling of CpG Islands in Human Breast Cancer Cells. *Hum Mol Genet* 8(3): 459-470.
- Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. & Rao, A. (2010). The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing. *PLoS ONE* 5(1): e8888.
- Irizarry, R. A., Ladd-Acosta, C., Carvalho, B., Wu, H., Brandenburg, S. A., Jeddeloh, J. A., Wen, B. & Feinberg, A. P. (2008). Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res* 18(5): 780-790.
- Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., Ji, H., Potash, J. B., Sabunciyan, S. & Feinberg, A. P. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 41(2): 178-186.
- Issa, J. P. (2004). CpG island methylator phenotype in cancer. *Nat Rev Cancer* 4(12): 988-993.
- Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466(7310): 1129-1133.
- Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. & Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333(6047): 1300-1303.
- Jaenisch, R. & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33 Suppl: 245-254.
- Jang, S. J., Soria, J. C., Wang, L., Hassan, K. A., Morice, R. C., Walsh, G. L., Hong, W. K. & Mao, L. (2001). Activation of melanoma antigen tumor antigens occurs early in lung carcinogenesis. *Cancer Res* 61(21): 7959-7963.
- Jenuwein, T. & Allis, C. D. (2001). Translating the histone code. *Science* 293(5532): 1074-1080.
- Ji, H., Ehrlich, L. I., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M. J., Irizarry, R. A., Kim, K., Rossi, D. J., Inlay, M. A., Serwold, T., Karsunky, H., Ho, L., Daley, G. Q., Weissman, I. L. & Feinberg, A. P. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* 467(7313): 338-342.
- Jiang, Y., Cui, L., Chen, W. D., Shen, S. H. & Ding, L. D. (2012). The prognostic role of RASSF1A promoter methylation in breast cancer: a meta-analysis of published data. *PLoS ONE* 7(5): e36780.
- Jin, S. G., Kadam, S. & Pfeifer, G. P. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* 38(11): e125.
- Johnson, T. B. & Coghill, R. D. (1925). RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS1. *Journal of the American Chemical Society* 47(11): 2838-2844.
- Jones, P. A. (1986). DNA methylation and cancer. *Cancer Res* 46(2): 461-466.
- Jones, P. A. & Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3(6): 415-428.
- Jones, P. A. & Baylin, S. B. (2007). The epigenomics of cancer. *Cell* 128(4): 683-692.
- Kaminsky, Z. A., Assadzadeh, A., Flanagan, J. & Petronis, A. (2005). Single nucleotide extension technology for quantitative site-specific evaluation of metC/C in GC-rich regions. *Nucleic Acids Res* 33(10): e95.

- Kantarjian, H. M., O'Brien, S., Cortes, J., Giles, F. J., Faderl, S., Issa, J. P., Garcia-Manero, G., Rios, M. B., Shan, J., Andreeff, M., Keating, M. & Talpaz, M. (2003). Results of decitabine (5-aza-2'deoxyctidine) therapy in 130 patients with chronic myelogenous leukemia. *Cancer* 98(3): 522-528.
- Karagiannis, T. C. & El-Osta, A. (2006a). Clinical potential of histone deacetylase inhibitors as stand alone therapeutics and in combination with other chemotherapeutics or radiotherapy for cancer. *Epigenetics* 1(3): 121-126.
- Karagiannis, T. C. & El-Osta, A. (2006b). Modulation of cellular radiation responses by histone deacetylase inhibitors. *Oncogene* 25(28): 3885-3893.
- Kelly, T. K., De Carvalho, D. D. & Jones, P. A. (2010). Epigenetic modifications as therapeutic targets. *Nat Biotech* 28(10): 1069-1078.
- Khulan, B., Thompson, R. F., Ye, K., Fazzari, M. J., Suzuki, M., Stasiek, E., Figueroa, M. E., Glass, J. L., Chen, Q., Montagna, C., Hatchwell, E., Selzer, R. R., Richmond, T. A., Green, R. D., Melnick, A. & Grealley, J. M. (2006). Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res* 16(8): 1046-1055.
- Killian, J. K., Bilke, S., Davis, S., Walker, R. L., Killian, M. S., Jaeger, E. B., Chen, Y., Hipp, J., Pittaluga, S., Raffeld, M., Cornelison, R., Smith, W. I., Jr., Bibikova, M., Fan, J. B., Emmert-Buck, M. R., Jaffe, E. S. & Meltzer, P. S. (2009). Large-scale profiling of archival lymph nodes reveals pervasive remodeling of the follicular lymphoma methylome. *Cancer Res* 69(3): 758-764.
- Kim, E. J., Lee, S. Y., Woo, M. K., Choi, S. I., Kim, T. R., Kim, M. J., Kim, K. C., Cho, E. W. & Kim, I. G. (2012). Fibulin-3 promoter methylation alters the invasive behavior of non-small cell lung cancer cell lines via MMP-7 and MMP-2 regulation. *Int J Oncol* 40(2): 402-408.
- Kim, M. S., Blake, M., Baek, J. H., Kohlhagen, G., Pommier, Y. & Carrier, F. (2003). Inhibition of histone deacetylase increases cytotoxicity to anticancer drugs targeting DNA. *Cancer Res* 63(21): 7291-7300.
- Kim, Y. H., Lee, H. C., Kim, S. Y., Yeom, Y. I., Ryu, K. J., Min, B. H., Kim, D. H., Son, H. J., Rhee, P. L., Kim, J. J., Rhee, J. C., Kim, H. C., Chun, H. K., Grady, W. M. & Kim, Y. S. (2011a). Epigenomic analysis of aberrantly methylated genes in colorectal cancer identifies genes commonly affected by epigenetic alterations. *Ann Surg Oncol* 18(8): 2338-2347.
- Kim, Y. J., Yoon, H. Y., Kim, S. K., Kim, Y. W., Kim, E. J., Kim, I. Y. & Kim, W. J. (2011b). EFEMP1 as a novel DNA methylation marker for prostate cancer: array-based DNA methylation and expression profiling. *Clin Cancer Res* 17(13): 4523-4530.
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994). HhaI methyltransferase flips its target base out of the DNA helix. *Cell* 76(2): 357-369.
- Klisovic, D. D., Katz, S. E., Effron, D., Klisovic, M. I., Wickham, J., Parthun, M. R., Guimond, M. & Marcucci, G. (2003). Depsipeptide (FR901228) inhibits proliferation and induces apoptosis in primary and metastatic human uveal melanoma cell lines. *Invest Ophthalmol Vis Sci* 44(6): 2390-2398.
- Knudson, A. G., Jr. (1970). Genetics and cancer. *Postgrad Med* 48(5): 70-74.
- Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R., Liu, X. S., Aravind, L., Agarwal, S., Maciejewski, J. P. & Rao, A. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 468(7325): 839-843.
- Kohli, R. M. & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502(7472): 472-479.
- Komashko, V. M. & Farnham, P. J. (2010). 5-azacytidine treatment reorganizes genomic histone modification patterns. *Epigenetics* 5(3).
- Kondo, T., Zhu, X., Asa, S. L. & Ezzat, S. (2007). The cancer/testis antigen melanoma-associated antigen-A3/A6 is a novel target of fibroblast growth factor receptor 2-IIIb through histone H3 modifications in thyroid cancer. *Clin Cancer Res* 13(16): 4713-4720.
- Kondo, Y., Shen, L., Cheng, A. S., Ahmed, S., Boumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B., Gold, D. L., Sekido, Y., Huang, T. H. & Issa, J. P. (2008). Gene

- silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat Genet* 40(6): 741-750.
- Korch, C. & Hagblom, P. (1986). In-vivo-modified gonococcal plasmid pJD1. A model system for analysis of restriction enzyme sensitivity to DNA modifications. *Eur J Biochem* 161(3): 519-524.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128(4): 693-705.
- Kriaucionis, S. & Bird, A. (2003). DNA methylation and Rett syndrome. *Hum Mol Genet* 12 Spec No 2: R221-227.
- Kriaucionis, S. & Heintz, N. (2009). The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* 324(5929): 929-930.
- Kristensen, L. S., Nielsen, H. M. & Hansen, L. L. (2009). Epigenetics and cancer treatment. *Eur J Pharmacol* 625(1-3): 131-142.
- Laird, P. W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* 11(3): 191-203.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., LeHoczek, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglu, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S. & Chen, Y. J. (2001). Initial sequencing and analysis of the human genome. *Nature* 409(6822): 860-921.

- Lapidus, R. G., Ferguson, A. T., Ottaviano, Y. L., Parl, F. F., Smith, H. S., Weitzman, S. A., Baylin, S. B., Issa, J. P. & Davidson, N. E. (1996). Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors. *Clin Cancer Res* 2(5): 805-810.
- Laurent, L., Wong, E., Li, G., Huynh, T., Tsigos, A., Ong, C. T., Low, H. M., Kin Sung, K. W., Rigoutsos, I., Loring, J. & Wei, C.-L. (2010). Dynamic changes in the human methylome during differentiation. *Genome Res*.
- Lecka-Czernik, B., Lumpkin, C. K., Jr. & Goldstein, S. (1995). An overexpressed gene transcript in senescent and quiescent human fibroblasts encoding a novel protein in the epidermal growth factor-like repeat family stimulates DNA synthesis. *Mol Cell Biol* 15(1): 120-128.
- Leonhardt, H., Page, A. W., Weier, H. U. & Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71(5): 865-873.
- Levy, A. (2008). Molecular and trophic mechanisms of tumorigenesis. *Endocrinol Metab Clin North Am* 37(1): 23-50, vii.
- Lewin, J., Schmitt, A. O., Adorjan, P., Hildmann, T. & Piepenbrock, C. (2004). Quantitative DNA methylation analysis based on four-dye trace data from direct sequencing of PCR amplicates. *Bioinformatics* 20(17): 3005-3012.
- Li, J. B., Gao, Y., Aach, J., Zhang, K., Kryukov, G. V., Xie, B., Ahlford, A., Yoon, J. K., Rosenbaum, A. M., Zaranek, A. W., LeProust, E., Sunyaev, S. R. & Church, G. M. (2009). Multiplex padlock targeted sequencing reveals human hypermutable CpG variations. *Genome Res* 19(9): 1606-1615.
- Li, L. C., Yeh, C. C., Nojima, D. & Dahiya, R. (2000). Cloning and characterization of human estrogen receptor beta promoter. *Biochem Biophys Res Commun* 275(2): 682-689.
- Lian, Christine G., Xu, Y., Ceol, C., Wu, F., Larson, A., Dresser, K., Xu, W., Tan, L., Hu, Y., Zhan, Q., Lee, C.-w., Hu, D., Lian, Bill Q., Kleffel, S., Yang, Y., Neiswender, J., Khorasani, Abraham J., Fang, R., Lezcano, C., Duncan, Lyn M., Scolyer, Richard A., Thompson, John F., Kakavand, H., Houvras, Y., Zon, Leonard I., Mihm Jr, Martin C., Kaiser, Ursula B., Schatton, T., Woda, Bruce A., Murphy, George F. & Shi, Yujiang G. (2012). Loss of 5-Hydroxymethylcytosine Is an Epigenetic Hallmark of Melanoma. *Cell* 150(6): 1135-1146.
- Liang, G., Gonzalgo, M. L., Salem, C. & Jones, P. A. (2002). Identification of DNA methylation differences during tumorigenesis by methylation-sensitive arbitrarily primed polymerase chain reaction. *Methods* 27(2): 150-155.
- Liang, G., Lin, J. C., Wei, V., Yoo, C., Cheng, J. C., Nguyen, C. T., Weisenberger, D. J., Egger, G., Takai, D., Gonzales, F. A. & Jones, P. A. (2004). Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci U S A* 101(19): 7357-7362.
- Lim, H. N. & van Oudenaarden, A. (2007). A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat Genet* 39(2): 269-275.
- Lima, S. C., Hernandez-Vargas, H., Simao, T., Durand, G., Kruel, C. D., Le Calvez-Kelm, F., Ribeiro Pinto, L. F. & Herceg, Z. (2011). Identification of a DNA methylome signature of esophageal squamous cell carcinoma and potential epigenetic biomarkers. *Epigenetics* 6(10): 1217-1227.
- Lindholm, J., Juul, S., Jorgensen, J. O., Astrup, J., Bjerre, P., Feldt-Rasmussen, U., Hagen, C., Jorgensen, J., Kosteljanetz, M., Kristensen, L., Laurberg, P., Schmidt, K. & Weeke, J. (2001). Incidence and late prognosis of cushing's syndrome: a population-based study. *J Clin Endocrinol Metab* 86(1): 117-123.
- Lippman, Z., Gendrel, A. V., Colot, V. & Martienssen, R. (2005). Profiling DNA methylation patterns using genomic tiling microarrays. *Nat Methods* 2(3): 219-224.
- Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B. & Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462(7271): 315-322.

- Liu, Y., Oakeley, E. J., Sun, L. & Jost, J. P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res* 26(4): 1038-1045.
- Mai, A. & Altucci, L. (2009). Epi-drugs to fight cancer: from chemistry to cancer treatment, the road ahead. *Int J Biochem Cell Biol* 41(1): 199-213.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. & Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057): 376-380.
- Martin-Subero, J. I. & Esteller, M. (2011). Profiling epigenetic alterations in disease. *Adv Exp Med Biol* 711: 162-177.
- Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S. & Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 33(18): 5868-5877.
- Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R. & Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437(7057): 436-439.
- Mill, J. & Petronis, A. (2009). Profiling DNA methylation from small amounts of genomic DNA starting material: efficient sodium bisulfite conversion and subsequent whole-genome amplification. *Methods Mol Biol* 507: 371-381.
- Mill, J., Yazdanpanah, S., Guckel, E., Ziegler, S., Kaminsky, Z. & Petronis, A. (2006). Whole genome amplification of sodium bisulfite-treated DNA allows the accurate estimate of methylated cytosine density in limited DNA resources. *Biotechniques* 41(5): 603-607.
- Mitra, S., Mazumder Indra, D., Basu, P. S., Mondal, R. K., Roy, A., Roychoudhury, S. & Panda, C. K. (2012). Alterations of RASSF1A in premalignant cervical lesions: clinical and prognostic significance. *Mol Carcinog* 51(9): 723-733.
- Mizukawa, B., Wei, J., Shrestha, M., Wunderlich, M., Chou, F. S., Griesinger, A., Harris, C. E., Kumar, A. R., Zheng, Y., Williams, D. A. & Mulloy, J. C. (2011). Inhibition of Rac GTPase signaling and downstream prosurvival Bcl-2 proteins as combination targeted therapy in MLL-AF9 leukemia. *Blood* 118(19): 5235-5245.
- Moran, A., Fernandez-Marcelo, T., Carro, J., De Juan, C., Pascua, I., Head, J., Gomez, A., Hernando, F., Torres, A. J., Benito, M. & Iniesta, P. (2012). Methylation profiling in non-small cell lung cancer: clinical implications. *Int J Oncol* 40(3): 739-746.
- Mottok, A., Renne, C., Willenbrock, K., Hansmann, M. L. & Brauner, A. (2007). Somatic hypermutation of SOCS1 in lymphocyte-predominant Hodgkin lymphoma is accompanied by high JAK2 expression and activation of STAT6. *Blood* 110(9): 3387-3390.
- Nakamura, N. & Takenaga, K. (1998). Hypomethylation of the metastasis-associated S100A4 gene correlates with gene activation in human colon adenocarcinoma cell lines. *Clin Exp Metastasis* 16(5): 471-479.
- Nomoto, S., Kanda, M., Okamura, Y., Nishikawa, Y., Qiyong, L., Fujii, T., Sugimoto, H., Takeda, S. & Nakao, A. (2010). Epidermal growth factor-containing fibulin-like extracellular matrix protein 1, EFEMP1, a novel tumor-suppressor gene detected in hepatocellular carcinoma using double combination array analysis. *Ann Surg Oncol* 17(3): 923-932.
- Obaya, A. J., Rua, S., Moncada-Pazos, A. & Cal, S. (2012). The dual role of fibulins in tumorigenesis. *Cancer Lett* 325(2): 132-138.

- Ogawa, O., Eccles, M. R., Szeto, J., McNoe, L. A., Yun, K., Maw, M. A., Smith, P. J. & Reeve, A. E. (1993). Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature* 362(6422): 749-751.
- Ogoshi, K., Hashimoto, S., Nakatani, Y., Qu, W., Oshima, K., Tokunaga, K., Sugano, S., Hattori, M., Morishita, S. & Matsushima, K. (2011). Genome-wide profiling of DNA methylation in human cancer cells. *Genomics* 98(4): 280-287.
- Ohtani-Fujita, N., Fujita, T., Aoike, A., Osifchin, N. E., Robbins, P. D. & Sakai, T. (1993). CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene* 8(4): 1063-1067.
- Okano, M., Bell, D. W., Haber, D. A. & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99(3): 247-257.
- Olek, A., Oswald, J. & Walter, J. (1996). A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res* 24(24): 5064-5066.
- Oshimo, Y., Nakayama, H., Ito, R., Kitadai, Y., Yoshida, K., Chayama, K. & Yasui, W. (2003). Promoter methylation of cyclin D2 gene in gastric carcinoma. *Int J Oncol* 23(6): 1663-1670.
- Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R., Lister, R., Ko, M., McLoughlin, E. M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G. Q., Liu, X. S., Ecker, J. R., Milos, P. M., Agarwal, S. & Rao, A. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473(7347): 394-397.
- Pietrobono, R., Pomponi, M. G., Tabolacci, E., Oostra, B., Chiurazzi, P. & Neri, G. (2002). Quantitative analysis of DNA demethylation and transcriptional reactivation of the FMR1 gene in fragile X cells treated with 5-azadeoxycytidine. *Nucleic Acids Research* 30(14): 3278-3285.
- Plimack, E. R., Kantarjian, H. M. & Issa, J. P. (2007). Decitabine and its role in the treatment of hematopoietic malignancies. *Leuk Lymphoma* 48(8): 1472-1481.
- Portela, A. & Esteller, M. (2010). Epigenetic modifications and human disease. *Nat Biotech* 28(10): 1057-1068.
- Pradhan, S., Bacolla, A., Wells, R. D. & Roberts, R. J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem* 274(46): 33002-33010.
- Qian, Z. R., Sano, T., Yoshimoto, K., Yamada, S., Ishizuka, A., Mizusawa, N., Horiguchi, H., Hirokawa, M. & Asa, S. L. (2005). Inactivation of RASSF1A tumor suppressor gene by aberrant promoter hypermethylation in human pituitary adenomas. *Lab Invest* 85(4): 464-473.
- Rakyan, V. K., Hildmann, T., Novik, K. L., Lewin, J., Tost, J., Cox, A. V., Andrews, T. D., Howe, K. L., Otto, T., Olek, A., Fischer, J., Gut, I. G., Berlin, K. & Beck, S. (2004). DNA Methylation Profiling of the Human Major Histocompatibility Complex: A Pilot Study for the Human Epigenome Project. *PLoS Biol* 2(12): e405.
- Rao, X., Evans, J., Chae, H., Pilrose, J., Kim, S., Yan, P., Huang, R. L., Lai, H. C., Lin, H., Liu, Y., Miller, D., Rhee, J. K., Huang, Y. W., Gu, F., Gray, J. W., Huang, T. M. & Nephew, K. P. (2012). CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene*.
- Rauch, T., Li, H., Wu, X. & Pfeifer, G. P. (2006). MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. *Cancer Res* 66(16): 7939-7947.
- Reik, W. & Dean, W. (2001). DNA methylation and mammalian epigenetics. *Electrophoresis* 22(14): 2838-2843.
- Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T. L., Wilson, C. J., Bell, S. P. & Young, R. A. (2000). Genome-wide location and function of DNA binding proteins. *Science* 290(5500): 2306-2309.

- Revell, K., Dudley, K. J., Clayton, R. N., McNicol, A. M. & Farrell, W. E. (2009). Loss of neuronatin expression is associated with promoter hypermethylation in pituitary adenoma. *Endocr Relat Cancer* 16(2): 537-548.
- Riggs, A. D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* 14(1): 9-25.
- Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L. & Wolffe, A. P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* 25(3): 338-342.
- Roessler, J., Ammerpohl, O., Gutwein, J., Hasemeier, B., Anwar, S. L., Kreipe, H. & Lehmann, U. (2012). Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. *BMC Res Notes* 5: 210.
- Roh, T. Y., Cuddapah, S., Cui, K. & Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A* 103(43): 15782-15787.
- Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlén, M. & Nyrén, P. (1996). Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry* 242(1): 84-89.
- Ronaghi, M., Uhlén, M., Nyrén, P. & Pål (1998). DNA SEQUENCING: A Sequencing Method Based on Real-Time Pyrophosphate. *Science* 281(5375): 363-365.
- Rountree, M. R., Bachman, K. E. & Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 25(3): 269-277.
- Rowther, F. B., Richardson, A., Clayton, R. N. & Farrell, W. E. (2010). Bromocriptine and dopamine mediate independent and synergistic apoptotic pathways in pituitary cells. *Neuroendocrinology* 91(3): 256-267.
- Sadr-Nabavi, A., Ramser, J., Volkmann, J., Naehrig, J., Wiesmann, F., Betz, B., Hellebrand, H., Engert, S., Seitz, S., Kreutzfeld, R., Sasaki, T., Arnold, N., Schmutzler, R., Kiechle, M., Niederacher, D., Harbeck, N., Dahl, E. & Meindl, A. (2009). Decreased expression of angiogenesis antagonist EFEMP1 in sporadic breast cancer is caused by aberrant promoter methylation and points to an impact of EFEMP1 as molecular biomarker. *Int J Cancer* 124(7): 1727-1735.
- Sakai, T., Toguchida, J., Ohtani, N., Yandell, D. W., Rapaport, J. M. & Dryja, T. P. (1991). Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet* 48(5): 880-888.
- Sasi, W., Jiang, W., Sharma, A. & Mokbel, K. (2010). Higher expression levels of SOCS 1,3,4,7 are associated with earlier tumour stage and better clinical outcome in human breast cancer. *BMC Cancer* 10(1): 178.
- Scheithauer, B. W., Gaffey, T. A., Lloyd, R. V., Sebo, T. J., Kovacs, K. T., Horvath, E., Yapici, O., Young, W. F., Jr., Meyer, F. B., Kuroki, T., Riehle, D. L. & Laws, E. R., Jr. (2006). Pathobiology of pituitary adenomas and carcinomas. *Neurosurgery* 59(2): 341-353; discussion 341-353.
- Scott, S. A., Lakshminikuttysamma, A., Sheridan, D. P., Sanche, S. E., Geyer, C. R. & DeCoteau, J. F. (2007). Zebularine inhibits human acute myeloid leukemia cell growth in vitro in association with p15INK4B demethylation and reexpression. *Exp Hematol* 35(2): 263-273.
- Seeliger, H., Camaj, P., Ischenko, I., Kleespies, A., De Toni, E. N., Thieme, S. E., Blum, H., Assmann, G., Jauch, K. W. & Bruns, C. J. (2009). EFEMP1 expression promotes in vivo tumor growth in human pancreatic adenocarcinoma. *Mol Cancer Res* 7(2): 189-198.
- Sharma, S., Kelly, T. K. & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis* 31(1): 27-36.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A. & Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119(7): 941-953.
- Simpson, D. J., Bicknell, J. E., McNicol, A. M., Clayton, R. N. & Farrell, W. E. (1999). Hypermethylation of the p16/CDKN2A/MTS1 gene and loss of protein expression is

- associated with nonfunctional pituitary adenomas but not somatotrophinomas. *Genes Chromosomes Cancer* 24(4): 328-336.
- Simpson, D. J., Clayton, R. N. & Farrell, W. E. (2002). Preferential loss of Death Associated Protein kinase expression in invasive pituitary tumours is associated with either CpG island methylation or homozygous deletion. *Oncogene* 21(8): 1217-1224.
- Simpson, D. J., Hibberts, N. A., McNicol, A. M., Clayton, R. N. & Farrell, W. E. (2000). Loss of pRb expression in pituitary adenomas is associated with methylation of the RB1 CpG island. *Cancer Res* 60(5): 1211-1216.
- Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J. & Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer Res* 63(18): 5821-5828.
- Smit, A. F. & Riggs, A. D. (1996). Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci U S A* 93(4): 1443-1448.
- Soengas, M. S., Capodici, P., Polsky, D., Mora, J., Esteller, M., Opitz-Araya, X., McCombie, R., Herman, J. G., Gerald, W. L., Lazebnik, Y. A., Cordon-Cardo, C. & Lowe, S. W. (2001). Inactivation of the apoptosis effector Apaf-1 in malignant melanoma. *Nature* 409(6817): 207-211.
- Song, C. X., Clark, T. A., Lu, X. Y., Kislyuk, A., Dai, Q., Turner, S. W., He, C. & Korlach, J. (2012). Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nat Methods* 9(1): 75-77.
- Song, C. X., Szulwach, K. E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C. H., Zhang, W., Jian, X., Wang, J., Zhang, L., Looney, T. J., Zhang, B., Godley, L. A., Hicks, L. M., Lahn, B. T., Jin, P. & He, C. (2011a). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol* 29(1): 68-72.
- Song, E. L., Hou, Y. P., Yu, S. P., Chen, S. G., Huang, J. T., Luo, T., Kong, L. P., Xu, J. & Wang, H. Q. (2011b). EFEMP1 expression promotes angiogenesis and accelerates the growth of cervical cancer in vivo. *Gynecol Oncol* 121(1): 174-180.
- Sutherland, E., Coe, L. & Raleigh, E. A. (1992). McrBC: a multisubunit GTP-dependent restriction endonuclease. *J Mol Biol* 225(2): 327-348.
- Suzuki, M. M. & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9(6): 465-476.
- Szwagierczak, A., Bultmann, S., Schmidt, C. S., Spada, F. & Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 38(19): e181.
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. & Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929): 930-935.
- Takai, D. & Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A* 99(6): 3740-3745.
- Teitz, T., Wei, T., Valentine, M. B., Vanin, E. F., Grenet, J., Valentine, V. A., Behm, F. G., Look, A. T., Lahti, J. M. & Kidd, V. J. (2000). Caspase 8 is deleted or silenced preferentially in childhood neuroblastomas with amplification of MYCN. *Nat Med* 6(5): 529-535.
- Uemura, H., Cho, M., Nakagawa, Y., Shimizu, K., Yoshikawa, M., Kim, S. & Hirao, Y. (2000). [MN/CA IX antigen as a potential target for renal cell carcinoma]. *Hinyokika Kiyo* 46(10): 745-748.
- Ushijima, T. (2005). Detection and interpretation of altered methylation patterns in cancer cells. *Nat Rev Cancer* 5(3): 223-231.
- Vaissiere, T., Cuenin, C., Paliwal, A., Vineis, P., Hoek, G., Krzyzanowski, M., Airoidi, L., Dunning, A., Garte, S., Hainaut, P., Malaveille, C., Overvad, K., Clavel-Chapelon, F., Linseisen, J., Boeing, H., Trichopoulou, A., Trichopoulos, D., Kaladidi, A., Palli, D., Krogh, V., Tumino, R., Panico, S., Bueno-De-Mesquita, H. B., Peeters, P. H., Kumle, M., Gonzalez, C. A., Martinez, C., Dorronsoro, M., Barricarte, A., Navarro, C., Quiros, J. R., Berglund, G., Janzon, L., Jarvholm, B., Day, N. E., Key, T. J., Saracci, R., Kaaks, R., Riboli, E., Hainaut, P. & Herceg, Z. (2009).

- Quantitative analysis of DNA methylation after whole bisulfite amplification of a minute amount of DNA from body fluids. *Epigenetics* 4(4): 221-230.
- Valk-Lingbeek, M. E., Bruggeman, S. W. & van Lohuizen, M. (2004). Stem cells and cancer; the polycomb connection. *Cell* 118(4): 409-418.
- Vandeva, S., Jaffrain-Rea, M. L., Daly, A. F., Tichomirowa, M., Zacharieva, S. & Beckers, A. (2010). The genetics of pituitary adenomas. *Best Pract Res Clin Endocrinol Metab* 24(3): 461-476.
- Wade, P. A. (2001). Methyl CpG binding proteins: coupling chromatin architecture to gene regulation. *Oncogene* 20(24): 3166-3173.
- Walter, K., Holcomb, T., Januario, T., Du, P., Evangelista, M., Kartha, N., Iniguez, L., Soriano, R., Huw, L., Stern, H., Modrusan, Z., Seshagiri, S., Hampton, G. M., Amler, L. C., Bourgon, R., Yauch, R. L. & Shames, D. S. (2012). DNA methylation profiling defines clinically relevant biological subsets of non-small cell lung cancer. *Clin Cancer Res* 18(8): 2360-2373.
- Walton, T. J., Li, G., Seth, R., McArdle, S. E., Bishop, M. C. & Rees, R. C. (2008). DNA demethylation and histone deacetylation inhibition co-operate to re-express estrogen receptor beta and induce apoptosis in prostate cancer cell-lines. *Prostate* 68(2): 210-222.
- Wang, T., Pan, Q., Lin, L., Szulwach, K. E., Song, C. X., He, C., Wu, H., Warren, S. T., Jin, P., Duan, R. & Li, X. (2012). Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum Mol Genet* 21(26): 5500-5510.
- Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. & Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37(8): 853-862.
- Wessely, F. & Emes, R. D. (2012). Identification of DNA methylation biomarkers from Infinium arrays. *Front Genet* 3: 161.
- Wilop, S., Fernandez, A. F., Jost, E., Herman, J. G., Brummendorf, T. H., Esteller, M. & Galm, O. (2011). Array-based DNA methylation profiling in acute myeloid leukaemia. *Br J Haematol* 155(1): 65-72.
- Woloschak, M., Yu, A. & Post, K. D. (1997). Frequent inactivation of the p16 gene in human pituitary tumors by gene methylation. *Mol Carcinog* 19(4): 221-224.
- Wu, H., Chen, Y., Liang, J., Shi, B., Wu, G., Zhang, Y., Wang, D., Li, R., Yi, X., Zhang, H., Sun, L. & Shang, Y. (2005). Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis. *Nature* 438(7070): 981-987.
- Wyatt, G. R. (1951). Recognition and estimation of 5-methylcytosine in nucleic acids. *Biochem J* 48(5): 581-584.
- Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., Huang, S., Min, J., Nicholson, T., Chen, T., Xu, G., Shi, Y., Zhang, K. & Shi, Yujiang G. (2011). Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Molecular cell* 42(4): 451-464.
- Yacqub-Usman, K., Duong, C. V., Clayton, R. N. & Farrell, W. E. (2012a). Epigenomic silencing of the BMP-4 gene in pituitary adenomas: a potential target for epidrug-induced re-expression. *Endocrinology* 153(8): 3603-3612.
- Yacqub-Usman, K., Richardson, A., Duong, C. V., Clayton, R. N. & Farrell, W. E. (2012b). The pituitary tumour epigenome: aberrations and prospects for targeted therapy. *Nat Rev Endocrinol* 8(8): 486-494.
- Yan, P. S., Perry, M. R., Laux, D. E., Asare, A. L., Caldwell, C. W. & Huang, T. H. (2000). CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. *Clin Cancer Res* 6(4): 1432-1438.
- Yang, A. S., Estecio, M. R., Garcia-Manero, G., Kantarjian, H. M. & Issa, J. P. (2003). Comment on "Chromosomal instability and tumors promoted by DNA hypomethylation" and "Induction of tumors in mice by genomic hypomethylation". *Science* 302(5648): 1153; author reply 1153.

- Yang, X., Ferguson, A. T., Nass, S. J., Phillips, D. L., Butash, K. A., Wang, S. M., Herman, J. G. & Davidson, N. E. (2000). Transcriptional activation of estrogen receptor alpha in human breast cancer cells by histone deacetylase inhibition. *Cancer Res* 60(24): 6890-6894.
- Yang, X., Lay, F., Han, H. & Jones, P. A. (2010). Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol Sci* 31(11): 536-546.
- Yang, X., Phillips, D. L., Ferguson, A. T., Nelson, W. G., Herman, J. G. & Davidson, N. E. (2001). Synergistic activation of functional estrogen receptor (ER)-alpha by DNA methyltransferase and histone deacetylase inhibition in human ER-alpha-negative breast cancer cells. *Cancer Res* 61(19): 7025-7029.
- Yoder, J. A., Soman, N. S., Verdine, G. L. & Bestor, T. H. (1997). DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J Mol Biol* 270(3): 385-395.
- Yoon, M. S., Suh, D. S., Choi, K. U., Sol, M. Y., Shin, D. H., Park, W. Y., Lee, J. H., Jeong, S. M., Kim, W. G. & Shin, N. R. (2010). High-throughput DNA hypermethylation profiling in different ovarian epithelial cancer subtypes using universal bead array. *Oncol Rep* 24(4): 917-925.
- Yoshida, M., Hoshikawa, Y., Koseki, K., Mori, K. & Beppu, T. (1990). Structural specificity for biological activity of trichostatin A, a specific inhibitor of mammalian cell cycle with potent differentiation-inducing activity in Friend leukemia cells. *J Antibiot (Tokyo)* 43(9): 1101-1106.
- Yu, F., Zingler, N., Schumann, G. & Strätling, W. H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic Acids Research* 29(21): 4493-4501.
- Yu, Y., Xu, F., Peng, H., Fang, X., Zhao, S., Li, Y., Cuevas, B., Kuo, W. L., Gray, J. W., Siciliano, M., Mills, G. B. & Bast, R. C., Jr. (1999). NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas. *Proc Natl Acad Sci U S A* 96(1): 214-219.
- Zeller, C., Dai, W., Steele, N. L., Siddiq, A., Walley, A. J., Wilhelm-Benartzi, C. S., Rizzo, S., van der Zee, A., Plumb, J. A. & Brown, R. (2012). Candidate DNA methylation drivers of acquired cisplatin resistance in ovarian cancer identified by methylome and expression profiling. *Oncogene* 31(42): 4567-4576.
- Zerbini, L. F. & Libermann, T. A. (2005). GADD45 deregulation in cancer: frequently methylated tumor suppressors and potential therapeutic targets. *Clin Cancer Res* 11(18): 6409-6413.
- Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W. & Arnheim, N. (1992). Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* 89(13): 5847-5851.
- Zhang, X., Sun, H., Danila, D. C., Johnson, S. R., Zhou, Y., Swearingen, B. & Klibanski, A. (2002). Loss of Expression of GADD45y, a Growth Inhibitory Gene, in Human Pituitary Adenomas: Implications for Tumorigenesis. *Journal of Clinical Endocrinology & Metabolism* 87(3): 1262-1267.
- Zhang, X., Zhou, Y., Mehta, K. R., Danila, D. C., Scolavino, S., Johnson, S. R. & Klibanski, A. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J Clin Endocrinol Metab* 88(11): 5119-5126.
- Zhao, J., Dahle, D., Zhou, Y., Zhang, X. & Klibanski, A. (2005). Hypermethylation of the promoter region is associated with the loss of MEG3 gene expression in human pituitary tumors. *J Clin Endocrinol Metab* 90(4): 2179-2186.
- Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., Ansell, P. J., Zhao, J., Weng, C. & Klibanski, A. (2007). Activation of p53 by MEG3 non-coding RNA. *J Biol Chem* 282(34): 24731-24742.
- Zhu, J. K. (2009). Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 43: 143-166.
- Zhu, W. G., Dai, Z., Ding, H., Srinivasan, K., Hall, J., Duan, W., Villalona-Calero, M. A., Plass, C. & Otterson, G. A. (2001). Increased expression of unmethylated CDKN2D by 5-aza-2'-deoxycytidine in human lung cancer cells. *Oncogene* 20(53): 7787-7796.

- Zhu, X., Asa, S. L. & Ezzat, S. (2007a). Ikaros is regulated through multiple histone modifications and deoxyribonucleic acid methylation in the pituitary. *Mol Endocrinol* 21(5): 1205-1215.
- Zhu, X., Asa, S. L. & Ezzat, S. (2008a). Fibroblast growth factor 2 and estrogen control the balance of histone 3 modifications targeting MAGE-A3 in pituitary neoplasia. *Clin Cancer Res* 14(7): 1984-1996.
- Zhu, X., Lee, K., Asa, S. L. & Ezzat, S. (2007b). Epigenetic silencing through DNA and histone methylation of fibroblast growth factor receptor 2 in neoplastic pituitary cells. *Am J Pathol* 170(5): 1618-1628.
- Zhu, X., Mao, X., Hurren, R., Schimmer, A. D., Ezzat, S. & Asa, S. L. (2008b). Deoxyribonucleic acid methyltransferase 3B promotes epigenetic silencing through histone 3 chromatin modifications in pituitary cells. *J Clin Endocrinol Metab* 93(9): 3610-3617.